APPLICATIONS AND GENERALIZATIONS OF THE CONCEPTION OF ADJOINT SYSTEMS*

BY

MAXIME BÔCHER

The highly fruitful conception of adjoint pairs of linear systems, each consisting of a homogeneous linear differential equation of the nth order and of n homogeneous linear boundary conditions, was first introduced in an explicit and general manner by BIRKHOFF.† That we have here a conception of fundamental importance is hardly yet generally appreciated. It is my purpose in the present paper to make an application of this idea to an important question in the theory of linear boundary problems which has so far been treated only in very special cases.‡ I will mention in passing that the results may also be obtained by the use of integral equations; though, if we wish to include all cases, not perhaps so simply or obviously as might at first sight appear.

In § 1 the conception of adjoint system and a few other fundamental matters concerning linear systems are explained and a few preliminary questions disposed of. In § 2 come the main results of the paper, namely Theorems I and II, which apply to the case of complex as well as to that of real independent variable. These theorems are then applied in some detail to the equation of the second order, whereby the relation to MASON'S results becomes apparent. In § 3 an extension to systems of equations of any order is briefly indicated; and a somewhat generalized type of differential equation is also considered. Finally, in § 4, the conception of adjoint system is extended to the case in which the number of boundary conditions is not the same as the order of the differential equation.

§ 1. *General Preliminaries.*

In the differential expression

$$L(u) = l_n \frac{d^n u}{dx^n} + l_{n-1} \frac{d^{n-1} u}{dx^{n-1}} + \cdots + l_0 u$$

* Presented to the Society April 26, 1913. Section 4 has been added since.

† These Transactions, vol. 9 (1908), p. 373.

‡ Cf. MASON, these Transactions, vol. 7 (1906), pp. 340–344, where the self-adjoint system of the second order is treated.

we assume that l_n does not vanish at any point of a certain finite region S which may be either a simply connected two-dimensional continuum in the complex x-plane, or any connected piece of the x-axis. In the former case we also assume that all the coefficients l_i are analytic functions of x throughout S; in the latter, that they are continuous but not necessarily real functions of the real variable x, and that l_i has continuous derivatives of the first i orders.*

This expression L is connected with its adjoint

$$M(v) = (-1)^n \frac{d^n (l_n v)}{dx^n} + (-1)^{n-1} \frac{d^{n-1} (l_{n-1} v)}{dx^{n-1}} + \cdots + l_0 v$$

by Lagrange's identity

$$vL(u) - uM(v) = \frac{d}{dx} P(u, v),$$

where P is a bilinear form in $u, u', \cdots u^{[n-1]}, v, v', \cdots v^{[n-1]}$. This bilinear form is non-singular for every value of x in S since in its matrix the elements in the secondary diagonal are alternately l_n and $-l_n$, while all the elements below this diagonal are zero.†

If we integrate this identity from a to b, these being any two points in S, we get Green's theorem, the second member of which is therefore a non-singular bilinear form in the quantities

$$(1) \qquad u(a), u'(a), \cdots u^{[n-1]}(a), u(b), u'(b), \cdots u^{[n-1]}(b),$$

$$(2) \qquad v(a), v'(a), \cdots v^{[n-1]}(a), v(b), v'(b), \cdots v^{[n-1]}(b).$$

Let $U_1(u), \cdots U_{2n}(u)$ be any $2n$ linearly independent linear forms in (1). The quantities (1) can then be expressed linearly in terms of $U_1, \cdots U_{2n}$. The bilinear form in the second member of Green's theorem thus becomes a linear form in the U's whose coefficients, which we denote by $V_{2n}(v), \cdots V_1(v)$, are linear forms in (2). Thus Green's theorem takes the form

$$(3) \qquad \int_a^b [vL(u) - uM(v)]\, dx = U_1(u) V_{2n}(v) + U_2(u) V_{2n-1}(v) + \cdots + U_{2n}(u) V_1(v).$$

* This restriction on the derivatives of the functions l_i is a sufficient but not a necessary one to secure the existence of an adjoint, that is of an expression $M(v)$ satisfying Lagrange's identity. For instance, if $n = 2$, it is sufficient to demand that l_2 have a continuous derivative l_2', and that $l_2' - l_1$ also have a continuous derivative. The adjoint is then

$$M(v) = l_2 \frac{d^2 v}{dx^2} + (2l_2' - l_1) \frac{dv}{dx} + \left[\frac{d}{dx}(l_2' - l_1) + l_0\right] v.$$

For still further generalization see the closing paragraph of § 3.

† This may be seen by a reference to the explicit form for P as given, for instance, in DARBOUX's *Théorie des Surfaces*, vol. 2, p. 100. It may also be readily inferred from Lagrange's identity itself if we merely admit the general form of this identity.

Moreover, since the bilinear form was non-singular, the V's must be linearly independent.

Now consider the linear system

$$L(u) = r, \tag{4}$$

$$U_i(u) = \gamma_i \qquad (i = 1, 2, \cdots n), \tag{5}$$

where L has the same meaning as before, r is a function of x continuous in S (and, if S is a two dimensional continuum in the complex plane, analytic in S), the γ_i's are constants, and the U_i's are linear forms in the quantities (1) which are restricted merely by being assumed linearly independent.

By the *reduced system* of (4), (5) we understand the homogeneous system

$$L(u) = 0, \tag{6}$$

$$U_i(u) = 0 \qquad (i = 1, 2, \cdots n). \tag{7}$$

Now form n further linear forms $U_{n+1}(u), \cdots U_{2n}(u)$ of the quantities (1) subject merely to the condition that $U_1, \cdots U_{2n}$ are linearly independent; and, by means of them, form, as above, the linear forms $V_1, \cdots V_{2n}$. The system

$$M(v) = 0, \tag{8}$$

$$V_i(v) = 0 \qquad (i = 1, 2, \cdots n) \tag{9}$$

is called a system adjoint to (6), (7). It is clear from the symmetry of the formulæ that, conversely, (6), (7) is a system adjoint to (8), (9)*.

A homogeneous system like (6), (7) is called compatible or incompatible according as it has or has not a solution other than zero. It is said to have k-fold compatibility if it has k and only k linearly independent solutions; and zero-fold compatibility shall be understood to mean incompatibility. If we indicate by $y_1, \cdots y_n$ linearly independent solutions of (6), we establish, by assuming the general solution of (6) in the form $c_1 y_1 + \cdots + c_n y_n$ and determining the c's so that this expression satisfies (7) also,

LEMMA I.† A necessary and sufficient condition that the system (6),

* It may be seen by very simple algebraic considerations that if we make a different choice of the arbitrary expressions $U_{n+1}, \cdots U_{2n}$, we shall obtain an adjoint system essentially identical with the one already found, the only difference being that the conditions (9) are r[illegible]ced by homogeneous linear combinations of themselves with non-vanishing determinant, and that $V_{n+1}, \cdots V_{2n}$ are changed only by having homogeneous linear combinations of $V_1, \cdots V_n$ added to them. These facts are not, however, necessary for what follows.

† Given substantially by MASON and BIRKHOFF, loc. cit.

(7) have k-fold compatibility is that the matrix

$$\begin{pmatrix} U_1(y_1) & \cdots & U_1(y_n) \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ U_n(y_1) & \cdots & U_n(y_n) \end{pmatrix} \tag{10}$$

be of rank $n-k$, where $y_1, \cdots y_n$ is a fundamental system of (6).

Similarly, if we denote by u_0 a particular solution of (4), so that the general solution is $u_0 + c_1 y_1 + \cdots + c_n y_n$, and if we determine these c's in such a way that this function satisfies (5) also, we find

LEMMA II. If the reduced system (6), (7) has k-fold compatibility, a necessary and sufficient condition that the complete system (4), (5) have a solution is that the matrix

$$\begin{pmatrix} U_1(y_1) & \cdots & U_1(y_n) & U_1(u_0) - \gamma_1 \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ U_n(y_1) & \cdots & U_n(y_n) & U_n(u_0) - \gamma_n \end{pmatrix}$$

be of rank $n-k$, where $y_1, \cdots y_n$ is a fundamental system of (6), and u_0 a particular solution of (4).

If $k = 0$, we see from Lemma I that the matrix last written is always of rank n. Hence

LEMMA III.* If the reduced system (6), (7) is incompatible, the complete system (4), (5) always has a solution.

We add finally

LEMMA IV. If $u_1, \cdots u_k$ are linearly independent solutions of the homogeneous system (6), (7), then the k sets of constants

$$U_{n+1}(u_i), \cdots U_{2n}(u_i) \qquad (i = 1, 2, \cdots k)$$

are linearly independent.

For otherwise there would exist k constants, $c_1, \cdots c_k$, not all zero, such that all the quantities

$$c_1 U_j(u_1) + \cdots + c_k U_j(u_k) \qquad (j = n+1, \cdots 2n)$$

vanish. Consequently if we let

$$u = c_1 u_1 + \cdots + c_k u_k,$$

we see, on the one hand, that u is not identically zero, and, on the other, that the equations

$$U_1(u) = 0, \cdots U_{2n}(u) = 0$$

* MASON, loc. cit., and, for a simpler proof, BÔCHER, Annals of Mathematics, 2d ser., vol. 13 (1911), p. 74. The proof here given is simpler than either of these.

are satisfied. These being $2n$ linearly independent, homogeneous, linear equations in the $2n$ quantities (1), it follows that all the quantities (1) vanish; but this is impossible since then u would vanish identically.

§ 2. *Conditions for Existence of Solutions of Non-Homogeneous Systems.*

THEOREM I.* *If a homogeneous system* (6), (7) *has k-fold compatibility, the adjoint system* (8), (9) *also has k-fold compatibility.*

On account of the reciprocal relation between a system and its adjoint, it will clearly be sufficient if we prove that (8), (9) has *at least* k-fold compatibility. For this purpose let u be any solution of the system (6), (7), and $z_1, \cdots z_n$, linearly independent solutions of the equation (8). From Green's theorem, (3), by substituting for v in succession the functons $z_1, \cdots z_n$, we find the equations

$$\begin{aligned} &U_{2n}(u)\,V_1(z_1) + \cdots + U_{n+1}(u)\,V_n(z_1) = 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &U_{2n}(u)\,V_1(z_n) + \cdots + U_{n+1}(u)\,V_n(z_n) = 0. \end{aligned} \tag{11}$$

These equations will be satisfied, in particular, if we replace u by any of the k linearly independent solutions $u_1, \cdots u_k$ of the system (6), (7). Consequently, regarded as a system of linear algebraic equations for determining $U_{2n}, \cdots U_{n+1}$, equations (11) have the k solutions

$$U_{2n}(u_i), \quad \cdots U_{n+1}(u_i) \qquad (i = 1, 2, \cdots k).$$

These solutions, by Lemma IV, are linearly independent. Consequently (11), regarded as a system of equations in the U's, has at least k linearly independent solutions. Hence, as we know from the theory of linear homogeneous algebraic equations, the matrix

$$\begin{pmatrix} V_1(z_1) & \cdots & V_n(z_1) \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ V_1(z_n) & \cdots & V_n(z_n) \end{pmatrix}$$

is of rank $n - k$ at most. From this, by Lemma I, we infer that the system (8), (9) has at least k-fold compatibility, as was to be proved.

We come now to our main result.

THEOREM II. *A necessary and sufficient condition that the complete system* (4), (5) *have a solution is that every solution, v, of the homogeneous adjoint system* (8), (9) *satisfy the relation*

$$\int_a^b v r\,dx = \gamma_1 V_{2n}(v) + \cdots + \gamma_n V_{n+1}(v). \tag{12}$$

* A special case of this theorem was established by BIRKHOFF, loc. cit.

In proving this theorem we may assume that the system (8), (9) has k-fold compatibility, where $k \geqq 1$; for if $k = 0$, the theorem to be proved follows at once from Lemma III and Theorem I.

If the system (4), (5) has a solution, and we substitute in Green's theorem, (3), for u this solution, and for v any solution of (8), (9), we see at once that the condition of Theorem II is a necessary condition. To prove it sufficient,* let v be any solution of the system (8), (9), and let u_0 be any solution of equation (4). From (3) we then get the equation

$$\int_a^b vr\,dx = U_1(u_0)\,V_{2n}(v) + \cdots + U_n(u_0)\,V_{n+1}(v).$$

By subtracting (12) from this equation, we find

$$(13)\quad (U_1(u_0) - \gamma_1)\,V_{2n}(v) + \cdots + (U_n(u_0) - \gamma_n)\,V_{n+1}(v) = 0.$$

Now let $y_1, \cdots y_n$ be linearly independent solutions of (6). Then from (3) we obtain the equations

$$(14)\qquad U_1(y_j)\,V_{2n}(v) + \cdots + U_n(y_j)\,V_{n+1}(v) = 0 \quad (j = 1, \cdots n).$$

Equations (13) and (14) together form a system of $n + 1$ linear homogeneous algebraic equations in $V_{2n}(v), \cdots V_{n+1}(v)$, so that if we denote by $v_1, \cdots v_k$ a system of linearly independent solutions of the system (8), (9), these linear algebraic equations have the k solutions

$$(15)\qquad V_{2n}(v_i), \quad \cdots V_{n+1}(v_i) \qquad (i = 1, 2, \cdots k),$$

which by Lemma IV are linearly independent. It follows from the theory of linear algebraic equations that the rank of the matrix of the system (13), (14) is at most $n - k$; and it cannot be of lower rank than this since the matrix (10) is of rank $n - k$. The matrix of the system (13), (14) being essentially the matrix of Lemma II, it follows that the system (4), (5) has a solution.

The condition of Theorem II is peculiarly simple in the important special case in which the γ_i's are all zero. If this is not so, we should be obliged to compute the quantities V_i before applying it. This computation we will now carry through in the case of the equation of the second order, which, without loss of generality, we take in the self-adjoint form

$$(16)\qquad L(u) \equiv \frac{d}{dx}\left(K\frac{du}{dx}\right) - Gu = r,$$

* This proof of sufficiency was worked out in the case $n = 2$ by Mr. Joseph Slepian, a student in one of my courses at Harvard.

where K does not vanish in S.* We then have, as may be easily verified

$$P(u, v) = K\left(v\frac{du}{dx} - u\frac{dv}{dx}\right).$$

The two boundary conditions (5) we will now write

$$\begin{aligned} a_1 u(a) + a_2 u(b) + a_3 u'(a) + a_4 u'(b) &= A, \\ b_1 u(a) + b_2 u(b) + b_3 u'(a) + b_4 u'(b) &= B. \end{aligned} \tag{17}$$

Since the first members of these equations have been assumed to be linearly independent, not all the determinants

$$d_{ij} = a_i b_j - a_j b_i$$

will be zero. We will prove the following facts:

If in Theorem II the differential equation is (16) *and the boundary conditions* (17), *the condition* (12) *may in general be written in any of the following forms*:

$$\begin{aligned} d_{12}\int_a^b vr\,dx &= K(a)(Ab_2 - Ba_2)v'(a) + K(b)(Ab_1 - Ba_1)v'(b), \\ d_{43}\int_a^b vr\,dx &= K(a)(Ab_4 - Ba_4)v(a) + K(b)(Ab_3 - Ba_3)v(b), \\ d_{23}\int_a^b vr\,dx &= K(a)(Ab_2 - Ba_2)v(a) - K(b)(Ab_3 - Ba_3)v'(b), \\ d_{14}\int_a^b vr\,dx &= K(a)(Ab_4 - Ba_4)v'(a) - K(b)(Ab_1 - Ba_1)v(b), \\ d_{13}\int_a^b vr\,dx &= K(a)(Ab_1 - Ba_1)v(a) + K(a)(Ab_3 - Ba_3)v'(a), \\ d_{42}\int_a^b vr\,dx &= K(b)(Ab_2 - Ba_2)v(b) + K(b)(Ab_4 - Ba_4)v'(b). \end{aligned} \tag{18}$$

Of these forms only those will be inadmissible for which the d_{ij} which occurs as a factor of the first member is zero.

In all cases, a necessary and sufficient condition that the system (16), (17) *be self-adjoint is*

$$d_{24}K(a) = d_{13}K(b). \tag{19}$$

* In order that, when x is restricted to real values, this expression, $L(u)$, satisfy the conditions stated in the text at the beginning of § 1, the function K must have continuous first and second derivatives. It will satisfy the conditions referred to in the first footnote of § 1 if K ha a continuous first derivative. The mere continuity of K, G, and r is, however, sufficient for our purposes as will be seen by a reference to the closing lines of § 3.

Let us write for brevity

$$u(a) = u_1, \quad u(b) = u_2, \quad u'(a) = u_3, \quad u'(b) = u_4,$$
$$v(a) = v_1, \quad v(b) = v_2, \quad v'(a) = v_3, \quad v'(b) = v_4.$$

The bilinear form in the second member of Green's theorem may now be written

$$K(b)(v_2 u_4 - u_2 v_4) - K(a)(v_1 u_3 - u_1 v_3), \tag{20}$$

and we have

$$U_1(u) = a_1 u_1 + a_2 u_2 + a_3 u_3 + a_4 u_4,$$
$$U_2(u) = b_1 u_1 + b_2 u_2 + b_3 u_3 + b_4 u_4.$$

Let us now first assume $d_{12} \neq 0$. In this case we may let

$$U_3 = u_3, \qquad U_4 = u_4,$$

and we readily compute the following expressions for the V's:

$$V_1 = \frac{1}{d_{12}}[K(b) d_{12} v_2 + K(a) d_{24} v_3 + K(b) d_{14} v_4],$$

$$V_2 = \frac{1}{d_{12}}[-K(a) d_{12} v_1 + K(a) d_{23} v_3 + K(b) d_{13} v_4],$$

$$V_3 = \frac{1}{d_{12}}[-K(a) a_2 v_3 - K(b) a_1 v_4],$$

$$V_4 = \frac{1}{d_{12}}[K(a) b_2 v_3 + K(b) b_1 v_4].$$

By equating the first two of these expressions to zero we get the adjoint boundary conditions, and we readily see that these are linear combinations of the original conditions (17) when and only when (19) is fulfilled. Thus we have established in this case that (19) is a necessary and sufficient condition that the system be self-adjoint.

On the other hand, by substituting in the condition of Theorem II the values just obtained for V_3, V_4 and making the necessary changes in notation, we find that this condition reduces to the first of conditions (18).

Assume next that $d_{13} \neq 0$. Letting

$$U_3 = u_2, \qquad U_4 = u_4,$$

we find

$$V_1 = \frac{1}{d_{13}}[K(a) d_{14} v_1 + K(b) d_{13} v_2 + K(a) d_{34} v_3],$$

$$V_2 = \frac{1}{d_{13}}[K(a)\, d_{12}\, v_1 - K(a)\, d_{23}\, v_3 - K(b)\, d_{13}\, v_4],$$

$$V_3 = \frac{1}{d_{13}}[-K(a)\, a_1\, v_1 - K(a)\, a_3\, v_3],$$

$$V_4 = \frac{1}{d_{13}}[K(a)\, b_1\, v_1 + K(a)\, b_3\, v_3].$$

Here again the first two expressions equated to zero give us the adjoint boundary conditions, from which it is readily inferred that (19) is a necessary and sufficient condition that the given system be self-adjoint. The expressions for V_3, V_4, on the other hand, show us that the condition of Theorem II reduces to the fifth of the relations (18).

We might proceed in a precisely similar way in the four other cases. This, however, is not necessary, for we may get the desired results by a mere interchange of letters and subscripts in those already obtained. Thus, to deduce the result in the case $d_{42} \neq 0$ from that last obtained when we assumed $d_{13} \neq 0$, we notice that the bilinear form (20) remains unchanged if we interchange the subscripts 1 with 2, and 3 with 4, and at the same time interchange $K(a)$ and $-K(b)$. Since these interchanges leave equation (19) unchanged, and carry over the fifth equation (18) into the sixth, our statement for the case $d_{24} \neq 0$ is established.*

Similarly, from what we have proved in the case $d_{12} \neq 0$, we can get the facts just stated for the case $d_{23} \neq 0$ by interchanging the subscripts 1 and 3 (leaving 2 and 4 unchanged) and changing $K(a)$ into $-K(a)$, these changes leaving the form (20) and the equation (19) unchanged, and carrying over the first equation (18) into the third. In a similar way we interchange the subscripts 2 and 4 and change the sign of $K(b)$ in order to get the results for the case $d_{14} \neq 0$ from those for $d_{12} \neq 0$; and also to get the results when $d_{43} \neq 0$ from those when $d_{23} \neq 0$.

The case treated by Mason is the self-adjoint case, in which (19) is satisfied.† Here, on account of the relation

$$d_{12}\, d_{34} + d_{13}\, d_{42} + d_{14}\, d_{23} = 0,$$

we see that, if the d_{ij}'s in the first four relations (18) are all zero, the product $d_{13}\, d_{42}$ is also zero, and hence, by (19), all the d_{ij}'s are zero, contrary to hypothesis. Consequently *in the self-adjoint case the condition of Theorem II may always be written in the form of at least one of the first four equations* (18).

*These substitutions also carry over the expressions found above for V_1, V_2, V_3, V_4 into new forms which may be used if $d_{42} \neq 0$. In doing this, the subscripts of the V's should be left unchanged. A similar remark applies to the next paragraph.

† MASON himself used the term self-adjoint here, guided by a somewhat vague sense of analogy, but without having in mind the conception of an adjoint system in other cases.

Without at first restricting ourselves to the self-adjoint case, let us suppose the reduced system has two-fold compatibility. Then every solution of the homogeneous equation $L(u) = 0$ satisfies the conditions:

$$a_1 u(a) + a_2 u(b) + a_3 u'(a) + a_4 u'(b) = 0,$$

$$b_1 u(a) + b_2 u(b) + b_3 u'(a) + b_4 u'(b) = 0.$$

From this we can easily infer that d_{13} and d_{42} cannot both be zero; for if they were, eliminating first $u'(b)$ and then $u(b)$ between the last written equations would give

$$d_{14} u(a) + d_{34} u'(a) = 0,$$

$$d_{12} u(a) + d_{32} u'(a) = 0,$$

and since these equations are to hold for all solutions u of the differential equation, for instance when $u(a) = 1$, $u'(a) = 0$, and also when $u(a) = 0$, $u'(a) = 1$, we should have

$$d_{12} = d_{43} = d_{23} = d_{14} = 0.$$

But this would make all the d_{ij}'s zero. Hence *if the reduced system is doubly compatible, the condition of Theorem II may always be written in the form of at least one of the last two equations* (18).

In the self-adjoint case, neither of the quantities d_{13}, d_{42} can be zero if the reduced system has two-fold compatibility, since by (19) they would both be zero, and this we have just seen is impossible. *If the reduced system is self-adjoint and doubly compatible, the condition of Theorem II can always be written in both of the last two forms given in* (18).

The results here stated, so far as they refer to the self-adjoint case, are essentially Mason's, though the statement here given is a trifle more complete than his.

§ 3. *Systems of Differential Equations.*

The extension to systems of linear differential equations is simple and fairly obvious, and will be here only briefly indicated.

Consider the differential expressions

$$L_{ij}(u) \equiv l_{ij}^{[n]} \frac{d^n u}{dx^n} + l_{ij}^{[n-1]} \frac{d^{n-1} u}{dx^{n-1}} + \cdots + l_{ij}^{[0]} u \quad \begin{pmatrix} i = 1, 2, \cdots m \\ j = 1, 2, \cdots m \end{pmatrix}$$

and their adjoints $M_{ij}(v)$. We assume that the coefficients of L_{ij} are not merely continuous throughout S, but are such that these adjoints exist and

have continuous coefficients.* We have then Lagrange's identity

$$vL_{ij}(u) - uM_{ij}(v) = \frac{d}{dx}P_{ij}(u, v). \tag{21}$$

Now consider the m homogeneous linear differential expressions of the mth order with the m dependent variables $u_1, \cdots u_m$:

$$L_i(u_1, \cdots u_m) \equiv \sum_{j=1}^{m} L_{ij}(u_j) \qquad (i = 1, 2, \cdots m). \tag{22}$$

By the adjoint of this system of expressions we understand the system

$$M_i(v_1, \cdots v_m) \equiv \sum_{j=1}^{m} M_{ji}(v_j) \qquad (i = 1, 2, \cdots m). \tag{23}$$

By replacing u by u_j and v by v_i in (21) and summing with regard to i and j, we find

$$\sum_{i=1}^{m}[v_i L_i(u_1, \cdots u_m) - u_i M_i(v_1, \cdots v_m)] = \frac{d}{dx}\sum_{i=1}^{m}\sum_{j=1}^{m} P_{ji}(u_i, v_j). \tag{24}$$

This we call Lagrange's identity for the system (22). The second member is the derivative of a bilinear form in $u_i, u_i', \cdots u_i^{[n-1]}, v_i, v_i', \cdots v_i^{[n-1]}$ $(i = 1, 2, \cdots m)$, whose determinant (cf. the second footnote to § 1) is, except perhaps for sign, the nth power of the determinant†

$$\Delta = \begin{vmatrix} l_{11}^{[n]} & \cdots & l_{1m}^{[n]} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ l_{m1}^{[n]} & \cdots & l_{mm}^{[n]} \end{vmatrix}.$$

We shall assume that Δ does not vanish at any point of S. Then integrating (24) from a to b, we get Green's theorem, in which the second member is a non-singular bilinear form in

$$u_i(a), \quad \cdots u_i^{[n-1]}(a), \quad u_i(b), \quad \cdots u_i^{[n-1]}(b) \tag{25}$$

$$v_i(a), \quad \cdots v_i^{[n-1]}(a), \quad v_i(b), \quad \cdots v_i^{[n-1]}(b) \quad (i = 1, \cdots m). \tag{26}$$

Let us now denote by $U_1(u_1, \cdots u_m), \cdots U_{2nm}(u_1, \cdots u_m)$ linearly independent linear forms in (25). We can then write Green's theorem in

* Cf. the first footnote to § 1. See also the closing lines of the present section.

† The proof of this statement requires a little manipulation of determinants, which, however, presents no difficulty.

the form

$$(27)\quad \int_a^b \sum_{i=1}^{m} [v_i L_i(u_1, \cdots u_m) - u_i M_i(v_1, \cdots v_m)]\, dx = U_1 V_{2nm} + \cdots + U_{2nm} V_1,$$

where the V_i's are linearly independent linear forms in (26).

Now consider the *complete system*

$$(28)\quad \begin{aligned} L_i(u_1, \cdots u_m) &= r_i \qquad (i = 1, \cdots m), \\ U_i(u_1, \cdots u_m) &= \gamma_i \qquad (i = 1, \cdots nm), \end{aligned}$$

the *reduced system*

$$(29)\quad \begin{aligned} L_i(u_1, \cdots u_m) &= 0 \qquad (i = 1, \cdots m), \\ U_i(u, \cdots u_m) &= 0 \qquad (i = 1, \cdots nm), \end{aligned}$$

and the *adjoint system*

$$(30)\quad \begin{aligned} M_i(v_1, \cdots v_m) &= 0 \qquad (i = 1, \cdots m), \\ V_i(v_1, \cdots v_m) &= 0 \qquad (i = 1, \cdots nm). \end{aligned}$$

The condition we have imposed, that Δ does not vanish in S, means precisely that the systems of differential equations in (28), (29), and (30) have no singular points in S. Consequently there exists one and only one system of functions which satisfy the differential equations of any one of these systems and at a point of S take on, along with their first $n - 1$ derivatives, arbitrarily given values.

By the order of compatibility of the system (29) we understand the number of linearly independent systems of functions $u_1, \cdots u_m$ which satisfy it. When we say that the system (29) is incompatible, or has zero-fold compatibility, we mean that the only systems of functions satisfying (29) are those which are all identically zero.

With these preliminaries, the extension of the methods and results of §§1, 2 to the case we are now considering presents no difficulty. We content ourselves, therefore, with merely stating the main results, which correspond to Theorems I and II.

THEOREM III. *If the homogeneous system* (29) *has k-fold compatibility, the adjoint system* (30) *also has k-fold compatibility.*

THEOREM IV. *A necessary and sufficient condition that the complete system* (28) *have a solution is that every solution* $(v_1, \cdots v_m)$ *of the homogeneous adjoint system* (30) *satisfy the relation*

$$\int_a^b [v_1 r_1 + \cdots + v_m r_m]\, dx = \gamma_1 V_{2nm}(v_1, \cdots v_m) + \cdots + \gamma_{nm} V_{nm+1}(v_1, \cdots v_m).$$

The case $n = 1$, in which the differential equations of the system are of the first order, is peculiarly important since, as is well known, all other cases may, if we wish, be reduced to this. We may here solve the equations for the derivatives, as we may always solve for the derivatives of highest order, and thus write

$$(31)\qquad \begin{aligned} L_i(u_1, \cdots u_m) &\equiv \frac{du_i}{dx} + a_{i1}(x)\,u_1 + \cdots + a_{im}(x)\,u_m \\ M_i(v_1, \cdots v_m) &\equiv -\frac{dv_i}{dx} + a_{1i}(x)\,v_1 + \cdots + a_{mi}(x)\,v_m \end{aligned} \qquad (i = 1, \cdots m).$$

and Green's Theorem becomes

$$(32)\qquad \int_a^b \sum_{i=1}^{m} [\,v_i L_i(u_1, \cdots u_m) - u_i M_i(v_1, \cdots v_m)\,]\,dx = -u_1(a)\,v_1(a) - \cdots - u_m(a)\,v_m(a) + u_1(b)\,v_1(b) + \cdots + u_m(b)\,v_m(b).$$

Consequently if

$$U_i(u_1, \cdots u_m) \equiv \alpha_{i1}u_1(a) + \cdots + \alpha_{im}u_m(a) + \beta_{i1}u_1(b) + \cdots + \beta_{im}u_m(b) \qquad (i = 1, \cdots 2m),$$

where the determinant D formed from the α's and β's is not zero, we shall have

$$V_{2m+1-i}(v_1, \cdots v_m) \equiv \frac{1}{D}[\,-A_{i1}\,v_1(a) - \cdots - A_{im}\,v_m(a) + B_{i1}\,v_1(b) + \cdots + B_{im}\,v_m(b)\,],$$

where A_{ij} and B_{ij} are the cofactors of α_{ij} and β_{ij} in the determinant D.

We now see that this case $n = 1$ is peculiarly simple in the following respect*: *In the case of a system of equations of the first order written in the form* (31) *the adjoint conditions* $V_i = 0$ *depend merely, and in a simple and explicit rational manner, on the original conditions* $U_i = 0$, *whereas in general, even in the case of a single differential equation of order higher than the first, they depend also on the coefficients of the differential equations.*

Another peculiarity worthy of remark in the case $n = 1$, when the form (31) is used, is that no demand on the coefficients of the differential equations need be made beyond their mere continuity in S, since the derivatives of these coefficients nowhere occur in our work, so that even the existence of their derivatives need not be demanded.

* This simple case has been considered by Bounitzky in Liouville's Journal, 6th series, vol. 5 (1909), p. 65, who, however, does not obtain the special case $n = 1$ of Theorems III and IV, being concerned with a quite different class of applications of adjoint systems.

By the aid of this last remark it is easy to see how all demands concerning the existence or continuity of derivatives of the coefficients of the differential equations may be avoided even when the order of these equations is greater than one: we have merely to replace these equations by systems of equations of the first order in the form (31) and apply Theorem IV to these systems. A little consideration, however, will show how we can accomplish the same object without reduction to a system of equations of the first order and can even generalize a step farther. For the sake of simplicity, I will explain this for the typical case of a single equation of the third order.

Let

$$L(u) \equiv a_3 \frac{d}{dx}\left\{ a_2 \frac{d}{dx}\left[a_1 \frac{d}{dx}(a_0 u) + b_1 u \right] + b_2 u \right\} + l_2 \frac{d}{dx}\left[a_1 \frac{d}{dx}(a_0 u) + b_1 u \right] + l_1 \frac{d}{dx}(a_0 u) + qu, \tag{33}$$

where the a's, b's, l's, and q are functions of x continuous throughout S. If these functions had continuous derivatives of certain orders, this would reduce at once to the form of expression considered in § 1. This will, in particular, always be the case if we are dealing with analytic functions. We wish, however, now to consider the case in which the a's, b's, l's, and q are continuous functions of the real variable x concerning which we make no assumption whatever as to the existence of derivatives. The expression $L(u)$ then includes the expressions of § 1 as very special cases.

The theory of the equation*

$$L(u) = r, \tag{34}$$

where r is also continuous in S, and where we assume that none of the functions a_0, a_1, a_2, a_3 vanishes in S, may be readily obtained by letting

$$y_1 = a_0 u, \qquad y_2 = b_1 u + a_1 y_1', \qquad y_3 = b_2 u + a_2 y_2'.$$

These functions, y_1, y_2, y_3, are then seen to satisfy a system of linear differential equations of the first order with no singular point in S, and we thus

* It should be noticed that this is not a differential equation in the ordinary sense since it is not a relation between u, u', u'', u'''. Indeed a solution of this equation will not in general possess even a first derivative. We shall, however, use the term *linear differential equation* in an extended sense to cover such equations as this. Their theory might readily be developed on an independent basis (instead of by the reduction to a system of the first order as indicated in the text), the method of successive approximations, for instance, being available to establish the existence theorem. The only mention of such equations which I have seen is the very special one by WEYL referred to at the close of this paper.

infer that the equation (34) has one and only one solution u for which

$$u, \qquad \frac{d}{dx}(a_0 u), \qquad \frac{d}{dx}\left[a_1 \frac{d}{dx}(a_0 u) + b_1 u\right] \tag{35}$$

take on arbitrarily given values at an arbitrarily given point of S. Moreover the general solution of (34) is readily seen to be obtained by adding to any particular solution the general solution of the reduced equation

$$L(u) = 0. \tag{36}$$

Finally, the solutions of this homogeneous equation are seen to form a ternary family, so that we may speak of any three linearly independent solutions as forming a fundamental system. In short, all the essential facts concerning linear differential equations apply without change to this case, with the exception of the change just noted in the existence theorem.

By the expression $M(v)$ adjoint to $L(u)$ we shall understand

$$\begin{aligned} M(v) = &- a_0 \frac{d}{dx}\left\{a_1 \frac{d}{dx}\left[a_2 \frac{d}{dx}(a_3 v) - l_2 v\right] + l_1 v\right\} \\ &+ b_1 \frac{d}{dx}\left[a_2 \frac{d}{dx}(a_3 v) - l_2 v\right] - b_2 \frac{d}{dx}(a_3 v) + qv. \end{aligned} \tag{37}$$

Between these adjoint expressions there is a Lagrange identity of precisely the form given in § 1, where, however, $P(u, v)$ now stands for a non-singular bilinear form not in u, u', u'', v, v', v'', but in the three quantities (35), as the first set of variables, and the three quantities

$$v, \qquad \frac{d}{dx}(a_3 v), \qquad \frac{d}{dx}\left[a_2 \frac{d}{dx}(a_3 v) - l_2 v\right] \tag{38}$$

as the second set. If we take for the U's linearly independent linear forms in the six values that (35) take on at a and b respectively, the V's, when we write Green's theorem in the form (3), will be linearly independent linear forms in the six values which the quantities (38) take on at the points a and b. With this understanding all the work in § 1 is applicable if L and M have the more general definitions (33) and (37), and Theorems I and II hold without further change for this more general case.* In particular, L may have the

* We might consider the even more general form of equation

$$\begin{aligned} a_2 \frac{d}{dx}\left\{a_2 \frac{d}{dx}\left[a_1 \frac{d}{dx}(a_0 u - r_0) + b_1 u - r_1\right] + b_2 u - r_2\right\} \\ + l_2 \frac{d}{dx}\left[a_1 \frac{d}{dx}(a_0 u - r_0) + b_1 u - r_1\right] + l_1 \frac{d}{dx}(a_0 u - r_0) + qu = r. \end{aligned}$$

form of § 1 where no restriction as to the differentiability of the coefficients is imposed. It should be noticed, however, that in this case M will no longer have the form of § 1, but another special case of the general form of this section.

The case in which L has the form (16), K being merely assumed continuous, is a very special case of what we have just been discussing. The fact that the existence of the derivative of K need not here be demanded has been mentioned by WEYL.*

§ 4. *An Extension of the Conception of Adjoint System.*

Let us now consider the system

$$L(u) = r \tag{4'}$$

$$U_i(u) = \gamma_i \qquad (i = 1, 2, \cdots m), \tag{5'}$$

where, instead of n boundary conditions, we have m. We shall assume these boundary conditions to be linearly independent, so that $m \gtreqless 2n$. The expression $L(u)$ may be taken either in the restricted sense of § 1 or in the broader sense explained at the end of § 3. By the side of this complete system, we have the reduced system

$$L(u) = 0 \tag{6'}$$

$$U_i(u) = 0 \qquad (i = 1, 2, \cdots m). \tag{7'}$$

We now form $2n - m$ further linear forms $U_{m+1}(u), \cdots U_{2n}(u)$ so that the functions $U_1, \cdots U_{2n}$ are linearly independent. From these, as in § 1, we form the expressions $V_1, \cdots V_{2n}$, which will also be linearly independent, such that Green's theorem may be written in the form (3). Then the system

$$M(v) = 0 \tag{8'}$$

$$V_i(v) = 0 \qquad (i = 1, 2, \cdots 2n - m) \tag{9'}$$

we call the system adjoint to (6'), (7').

This can readily be reduced to a system of linear differential equations of the first order, and we thus see that, apart from a certain obvious change in the existence theorem, the theory of this equation does not differ substantially from that of the ordinary non-homogeneous equation so far as the more fundamental theorems go. Nevertheless, since this equation cannot be written $L(u) = f(x)$, where L is homogeneous, the reasoning used to establish Theorem II is no longer applicable, and it is readily seen that this theorem must be replaced by a more complex one similar to (and of course a special case of) Theorem IV. The formula involved contains the integral of a sum of products, not that of a single product.

*Mathematische Annalen, vol. 68 (1910), p. 221.

If the three systems of § 1 are replaced by these three, Lemmas I, II, IV of that section hold with only the obvious modification that the matrices in Lemmas I and II now have m rows instead of n, while in Lemma IV the subscripts of the U's must begin with $m+1$. The proofs remain essentially unchanged.

Lemma III must be replaced by the following:

LEMMA III'. When $m \leqq n$, the complete system (4'), (5') always has a solution if the reduced system has $(n-m)$-fold compatibility.

The proof is substantially that of Lemma III.

A modification is also necessary in the statement of Theorem I.

THEOREM I'. *If a homogeneous system which involves m linearly independent boundary conditions, and where the differential equation is of order n, has k-fold compatibility, then $k+m-n \geqq 0$, and this number is precisely the order of compatibility of the adjoint system* (8'), (9').

The proof follows closely that given above for Theorem I. We show first that, if κ is the order of compatibility of (8'), (9'), $\kappa \geqq k+m-n$. Consequently, since (6'), (7') is the adjoint of (8'), (9'), we must also have $k \geqq \kappa + (2n-m) - n$; that is $\kappa \leqq k+m-n$. From a comparison of these two inequalities our theorem follows.

THEOREM II'. *A necessary and sufficient condition that the complete system* (4'), (5') *have a solution is that every solution, v, of the homogeneous adjoint system* (8'), (9') *satisfy the relation*

$$\int_a^b v r\,dx = \gamma_1 V_{2n}(v) + \cdots + \gamma_m V_{2n-m+1}(v).$$

Here again the proof is substantially that of Theorem II, the case $n \geqq m$, $k = n-m$ being first disposed of by means of Lemma III'.

This theorem of course contains Theorem II, and hence Mason's results, as a special case. The only other special case I have found in the literature is that in which $m = n+1$ (n arbitrary), $U_1(u) = u(a)$, $U_2(u) = u'(a)$, $\cdots U_n(u) = u^{[n-1]}(a)$, $U_{n+1}(u) = \beta u(b) + \beta' u'(b) + \cdots + \beta^{[n-1]} u^{[n-1]}(b)$; while $\gamma_1 = \cdots = \gamma_n = 0$. In this case, $k = 0$, $\kappa = 1$. There is therefore essentially only one solution, v, of the adjoint system, and the condition that the complete system have a solution is

$$\int_a^b v r\,dx = 0.$$

A condition of exactly this form is obtained by DINI* by quite another

*Annali di Matematica, vol. 12 (1906), p. 226.

method, and it is not obvious at first sight that his function v (called in his notation $D_{1n}/a_0 D$) is the same as the one our method gives. This is, however, actually the case, as may be easily proved, and consequently this result of Dini may be regarded as a very special case, stated in a somewhat different form, of Theorem II'.

The possibility of extending the results of this section to systems of differential equations is obvious.

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

ON A CERTAIN CLASS OF SELF-PROJECTIVE SURFACES*

BY

E. J. WILCZYNSKI

Introduction.

Let P be a point of a surface S, and let R be a non-vanishing two-dimensional portion of S, finite or infinite in extent, such that P is one of its interior points and such that the homogeneous coördinates $x_1, \cdots, x_4$ of any point of R may be expressed as convergent series of positive integral powers of two parameters u and v. Moreover, let the two asymptotic tangents of the surface at P be distinct, and let us assume further that neither of these asymptotic tangents has contact of higher than the second order with the surface at P. The tetrahedron of reference may then be chosen in such a way that the surface may be represented by a canonical development of the form

$$z = xy + \frac{1}{6}(x^3 + y^3) + \frac{1}{24}(Ix^4 + Jy^4) + \cdots,$$

where the non-homogeneous coördinates are defined by the equations

$$x = \frac{x_1}{x_4}, \qquad y = \frac{x_2}{x_4}, \qquad z = \frac{x_3}{x_4},$$

and where I, J and all further coefficients of the development are absolute differential invariants of the surface.†

It is the purpose of this paper to study those surfaces for which the two invariants I and J are equal to zero, at all points which satisfy the conditions formulated above.

§ 1. *Partial differential equations of the surfaces referred to their asymptotic curves.*

If the curves $u =$ const. and $v =$ const. of the surface

$$y^{(k)} = f^{(k)}(u, v) \qquad (k = 1, 2, 3, 4)$$

* Presented to the Society, September 11, 1912.

† E. J. WILCZYNSKI, *Projective Differential Geometry of Curved Surfaces (Second Memoir)*. These Transactions, vol. 9 (1908), p. 103. We shall hereafter refer to this paper as Second Memoir.

are asymptotic curves, there will exist a completely integrable system of partial differential equations*

$$\begin{aligned} y_{uu} + 2by_v + fy &= 0, \\ y_{vv} + 2a'y_u + gy &= 0, \end{aligned} \tag{1}$$

of which $y^{(1)}, \cdots, y^{(4)}$ are four linearly independent solutions. The integrability conditions of this system are†

$$\begin{aligned} a'_{uu} + g_u + 2ba'_v + 4a'b_v &= 0, \\ b_{vv} + f_v + 2a'b_u + 4ba'_u &= 0, \\ g_{uu} - f_{vv} - 4fa'_u - 2a'f_u + 4gb_v + 2bg_v &= 0. \end{aligned} \tag{2}$$

Let us put

$$A = a'b^2, \qquad B = a'^2 b. \tag{3}$$

Then the invariants I and J, which occur in the canonical development, have the following values‡

$$I = \frac{B_u}{4B\sqrt[3]{A}}, \qquad J = \frac{A_v}{4A\sqrt[3]{B}},$$

where, under our assumptions, A and B are both different from zero. For if either a' or b were equal to zero identically the surface S would be a ruled surface, and one of the asymptotic tangents of such a point would have higher than second order contact with the surface. But this is a case which we have explicitly excluded from consideration.

If I and J are identically equal to zero, we must have

$$A = U, \qquad B = V,$$

where U and V are non-vanishing functions of the single variables u and v respectively. Now A and B are relative invariants of the system of differential equations (1). If we introduce new independent variables $\bar{u}$ and $\bar{v}$ by means of the transformation

$$\bar{u} = \alpha(u), \qquad \bar{v} = \beta(v),$$

the new values $\bar{A}$ and $\bar{B}$ of these invariants will be connected with their original values A and B by the following relations§,

$$\bar{A} = \frac{A}{\alpha_u^3}, \qquad \bar{B} = \frac{B}{\beta_v^3}. \tag{4}$$

* *First Memoir*, these Transactions, vol. 8 (1907), p. 246.
† *Loc. cit.*, p. 247.
‡ *Second Memoir*, p. 103.
§ *First Memoir*, p. 249.

Since in our case A is a function of u alone, and B a function of v alone, the functions $\alpha(u)$ and $\beta(v)$ may be chosen in such a way as to reduce $\bar{A}$ and $\bar{B}$ to unity. Let us assume that this reduction has been made, so that

$$A = B = 1.$$

We conclude from (3) that

$$a' = b = 1 \quad \text{or} \quad \omega \quad \text{or} \quad \omega^2,$$

where ω is an imaginary cube root of unity. But we may, without any restriction of generality, assume

$$a' = b = 1.$$

For, as we see from (4), the conditions $A = B = 1$ will not be disturbed by a transformation of the independent variables for which $\alpha_u^3 = \beta_v^3 = 1$. Now a' and b are themselves relative invariants and, if their common value is ω or ω^2, it is possible to find a linear transformation of the independent variables reducing both of them to unity.

If we substitute $a' = b = 1$ into the integrability conditions (2), they reduce to

$$g_u = 0, \qquad f_v = 0, \qquad -f_u + g_v = 0,$$

so that

$$f = U, \qquad g = V,$$

U and V being functions of the single variables u and v respectively, and

$$U_u = V_v.$$

Since the left member is a function of u alone, and the right member depends only upon v, their common value must be a constant c_0, so that finally

$$f = c_0 u + c_1, \qquad g = c_0 v + c_2.$$

We have found the following theorem. *The integral surfaces of the completely integrable system of partial differential equations*

$$\begin{aligned} &y_{uu} + 2y_v + (c_0 u + c_1)\, y = 0, \\ &y_{vv} + 2y_u + (c_0 v + c_2)\, y = 0 \end{aligned} \tag{5}$$

are the only surfaces at all of whose points the invariants I and J are equal to zero.

These surfaces have the following characteristic property. *The canonical cubic of every point P of such a surface S has fourth order contact with S at P.**

* In these and all similar statements, the locution "every point" is intended to mean every point which satisfies the conditions formulated at the beginning of the introduction. For the definition of the "canonical cubic," cf. *Second Memoir*, p. 108.

§ 2. *Determination of the surfaces for the case when $c_0 = 0$ and when the characteristic equations have distinct roots.**

If $c_0 = 0$, system (5) reduces to

$$\begin{aligned} y_{uu} + 2y_v + c_1 y &= 0, \\ y_{vv} + 2y_u + c_2 y &= 0, \end{aligned} \tag{6}$$

and may be integrated by elementary functions. In fact we may put

$$y = e^{\alpha u + \beta v},$$

where α and β are constants which satisfy the *characteristic* equations

$$\begin{aligned} \alpha^2 + 2\beta + c_1 &= 0, \\ \beta^2 + 2\alpha + c_2 &= 0. \end{aligned} \tag{7}$$

Elimination of β gives

$$\alpha^4 + 2c_1\alpha^2 + 8\alpha + c_1^2 + 4c_2 = 0, \qquad \beta = -\tfrac{1}{2}(c_1 + \alpha^2), \tag{8}$$

and similarly we find

$$\beta^4 + 2c_2\beta^2 + 8\beta + c_2^2 + 4c_1 = 0, \qquad \alpha = -\tfrac{1}{2}(c_2 + \beta^2). \tag{9}$$

The discriminants of (8) and (9), except for a numerical factor, are both equal to

$$\Delta = c_1^2 c_2^2 + 4(c_1^3 + c_2^3) + 18c_1 c_2 - 27. \tag{10}$$

We shall first assume that Δ is not equal to zero, so that the four roots $\alpha_1, \cdots, \alpha_4$ of (8) and $\beta_1, \cdots, \beta_4$ of (9) are distinct. We shall then have four solutions of the form

$$y_k = e^{\alpha_k u + \beta_k v} \qquad (k = 1, 2, 3, 4), \tag{11}$$

which are linearly independent. The surface may therefore be represented by equations (11), the curves $u =$ const. and $v =$ const. being its asymptotic lines.

We proceed to find the equation of the surface by eliminating u and v. We find

$$\log \frac{y_2^{\alpha_3-\alpha_1} y_1^{\alpha_2-\alpha_3}}{y_3^{\alpha_2-\alpha_1}} = Av, \qquad \log \frac{y_2^{\alpha_4-\alpha_1} y_1^{\alpha_2-\alpha_4}}{y_4^{\alpha_2-\alpha_1}} = Bv, \tag{12}$$

where

$$\begin{aligned} A &= \beta_1(\alpha_2 - \alpha_3) + \beta_2(\alpha_3 - \alpha_1) + \beta_3(\alpha_1 - \alpha_2) \\ &= \tfrac{1}{2}(\alpha_2 - \alpha_3)(\alpha_3 - \alpha_1)(\alpha_1 - \alpha_2), \\ B &= \beta_1(\alpha_2 - \alpha_4) + \beta_2(\alpha_4 - \alpha_1) + \beta_4(\alpha_1 - \alpha_2) \\ &= \tfrac{1}{2}(\alpha_2 - \alpha_4)(\alpha_4 - \alpha_1)(\alpha_1 - \alpha_2). \end{aligned}$$

* Discussion of the cases when $c_0 \neq 0$ will be reserved for a subsequent paper.

Consequently, if we introduce non-homogeneous coördinates by putting

$$X = \frac{y_2}{y_1}, \qquad Y = \frac{y_3}{y_1}, \qquad Z = \frac{y_4}{y_1},$$

and if we use the notation

$$\lambda = \frac{\alpha_3 - \alpha_1}{\alpha_2 - \alpha_1}, \qquad \mu = \frac{\alpha_4 - \alpha_1}{\alpha_2 - \alpha_1},$$

we find

$$(X^\lambda Y^{-1})^{\mu(1-\mu)} = (X^\mu Z^{-1})^{\lambda(1-\lambda)}$$

or

$$X = Y^{\frac{1-\mu}{\lambda(\lambda-\mu)}} Z^{-\frac{1-\lambda}{\mu(\lambda-\mu)}} = Y^\alpha Z^\beta \tag{13}$$

as the equation of our surface, which is clearly self-projective, being left invariant by the transformations of a two-parameter projective group.

The quantities λ and μ, and also α and β, are clearly functions of c_1 and c_2. The question immediately suggests itself whether λ and μ may be chosen as arbitrary constants or whether they are subject to some restrictions. It is clear in the first place that we must assume

$$\lambda \neq 0, 1, \infty; \quad \mu \neq 0, 1, \infty, \quad \lambda \neq \mu,$$

since the four roots $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ of (8) are supposed to be distinct. With this understanding we find

$$\alpha_3 = \alpha_1 + \lambda(\alpha_2 - \alpha_1), \quad \alpha_4 = \alpha_1 + \mu(\alpha_2 - \alpha_1), \quad \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 0.$$

Of course, one of the roots of (8) may be equal to zero. But, since we are now dealing with the case of distinct roots, there can be only one such root. Let the notation be fixed so that $\alpha_1 \neq 0$, and therefore $\lambda + \mu + 1 \neq 0$. We shall find

$$\alpha_2 = \frac{\lambda + \mu - 3}{\lambda + \mu + 1}\alpha_1, \qquad \alpha_3 = \frac{-3\lambda + \mu + 1}{\lambda + \mu + 1}\alpha_1, \qquad \alpha_4 = \frac{\lambda - 3\mu + 1}{\lambda + \mu + 1}\alpha_1.$$

According to (8) the sum of the products of $\alpha_i, \alpha_j, \alpha_k$, taken three at a time, must be equal to -8. Thus

$$\begin{aligned} \alpha_1^3 [& (\lambda + \mu - 3)(-3\lambda + \mu + 1)(\lambda - 3\mu + 1) \\ & + (-3\lambda + \mu + 1)(\lambda - 3\mu + 1)(\lambda + \mu + 1) \\ & + (\lambda - 3\mu + 1)(\lambda + \mu - 3)(\lambda + \mu + 1) \\ & + (\lambda + \mu - 3)(-3\lambda + \mu + 1)(\lambda + \mu + 1)] \\ & = -8(\lambda + \mu + 1)^3. \end{aligned} \tag{14}$$

If the coefficient of α_1^3 in this equation is not equal to zero, we can therefore

express $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ as functions of λ and μ. The two remaining symmetric functions of the roots of (8) will then give us $2c_1$ and $c_1^2 + 4c_2$, so that c_1 and c_2 are expressible as algebraic functions of λ and μ. In other words; if λ and μ are any two numbers which do not cause the above expression to vanish, the surface (13) will be an integral surface of a system of differential equations of form (6).

The coefficient of α_1^3 in (14) may be reduced to -8 times the expression

$$\lambda^3 + \mu^3 - \lambda^2\mu - \lambda\mu^2 - \lambda^2 - \mu^2 + 2\lambda\mu - \lambda - \mu + 1$$

and this vanishes if and only if the anharmonic curve $v = \text{const.}$, represented by equations (12), belongs to a linear complex.* But surfaces whose asymptotic lines belong to linear complexes are indeed excluded from our class of surfaces. In fact the differential equation of the fourth order characteristic of the asymptotic lines $v = \text{const.}$ is found to be

$$y_{uuuu} + 2c_1 y_{uu} + 8y_u + (c_1^2 + 4c_2)\, y = 0,$$

whose invariant of weight three is found to be independent of c_1 and c_2, namely equal to 2. In no case then, can this invariant vanish, i. e., *in no case will an asymptotic curve of the surface S belong to a linear complex. The asymptotic curves of S are clearly anharmonic curves whose invariants are subject to no other restriction than this, the curves of each family being projectively equivalent to each other, but in general (if $c_1 \neq c_2$) not equivalent to those of the other family.* Moreover, the invariants of each family determine those of the other and of the surface.†

Thus, *any surface of the form* (13) *may be regarded as an integral surface of system* (6), *i. e., as a surface for which I and J are everywhere equal to zero, provided that λ and μ are different from each other, from* 0, 1, *or* ∞, *and do not satisfy the condition*

$$\lambda^3 + \mu^3 - \lambda^2\mu - \lambda\mu^2 - \lambda^2 - \mu^2 + 2\lambda\mu - \lambda - \mu + 1 = 0. \tag{15}$$

The corresponding restrictions on the exponents α and β, which enter into the equation of the surface, may be found as follows. We have

$$\alpha = \frac{1-\mu}{\lambda(\lambda-\mu)}, \qquad \beta = -\frac{1-\lambda}{\mu(\lambda-\mu)},$$

whence

$$-\alpha(\alpha+\beta)\lambda^2 + 2\alpha\lambda + \beta - 1 = 0, \qquad \mu = \frac{1-\alpha\lambda^2}{1-\alpha\lambda}, \tag{16a}$$

* E. J. WILCZYNSKI, *Projective Differential Geometry of Curves and Ruled Surfaces*, B. G. Teubner, Leipzig, 1906, p. 281. In the notation there used $\theta_3 = 0$ is the condition for a curve belonging to a linear complex. We shall refer to this book as Proj. Diff. Geom.

† *Loc. cit.*, p. 281.

and

$$(16b) \qquad -\beta(\alpha+\beta)\mu^2+2\beta\mu+\alpha-1=0, \quad \lambda=\frac{1-\beta\mu^2}{1-\beta\mu}.$$

Since α and β are supposed to be finite, and λ and μ must not be equal to $0, 1$ or ∞, we must have

$$\alpha\neq 1, \quad \beta\neq 1, \quad -\alpha(\alpha+\beta)+2\alpha+\beta-1\neq 0, \quad \alpha(\alpha+\beta)\neq 0,$$
$$-\beta(\alpha+\beta)+2\beta+\alpha-1\neq 0, \quad \beta(\alpha+\beta)\neq 0.$$

But, if $\lambda=1$, μ will also be equal to unity. Therefore the third and fourth inequalities must hold simultaneously. But we may write these inequalities as follows

$$(\alpha-1)(\alpha+\beta-1)\neq 0, \qquad (\beta-1)(\alpha+\beta-1)\neq 0,$$

whence $\alpha+\beta-1\neq 0$. We must also exclude the possibility $\lambda=\mu$. But, if $\lambda=\mu$, we find

$$\lambda-\alpha\lambda^2=1-\alpha\lambda^2,$$

so that λ is either equal to one or infinity, which cases have already been discussed. We see therefore that α and β are subject to the inequalities

$$(17) \qquad \alpha\neq 0, \quad \alpha\neq 1, \quad \beta\neq 0, \quad \beta\neq 1, \quad \alpha+\beta\neq 0, \quad \alpha+\beta\neq 1.$$

From (16) we find

$$\lambda=\frac{\alpha\beta\pm\beta t}{\alpha\beta(\alpha+\beta)}, \qquad \mu=\frac{\alpha\beta\pm\alpha t}{\alpha\beta(\alpha+\beta)}, \qquad t^2=\alpha\beta(\alpha+\beta-1),$$

and therefore

$$\lambda+\mu=\frac{2\alpha\beta\pm(\beta-\alpha)t}{\alpha\beta(\alpha+\beta)}, \qquad \lambda\mu=\frac{\alpha\beta(2-\alpha-\beta)\pm(\beta-\alpha)t}{\alpha\beta(\alpha+\beta)^2},$$

$$(\lambda-\mu)^2=\frac{(\alpha+\beta-1)^2}{\alpha\beta},$$

$$\lambda\mu(\lambda+\mu)=\frac{2\alpha\beta(2-\alpha-\beta)+(\beta-\alpha)^2(\alpha+\beta-1)\pm(\beta-\alpha)(4-\alpha-\beta)t}{\alpha\beta(\alpha+\beta)^3},$$

$$(\lambda+\mu)^3=\frac{\alpha\beta[8\alpha\beta+6(\beta-\alpha)^2(\alpha+\beta-1)]\pm(\beta-\alpha)[12\alpha\beta+(\beta-\alpha)^2(\alpha+\beta-1)]t}{\alpha^2\beta^2(\alpha+\beta)^3}.$$

Combined with these equations, the final restriction that λ and μ shall not satisfy (15) gives rise to the inequality

$$(18) \qquad \alpha\beta[(\alpha+\beta)(\alpha+\beta-2)(\alpha-1)(\beta-1)-4\alpha\beta]^2 \neq(\alpha+\beta)^2(\alpha-\beta)^2(\alpha-1)^2(\beta-1)^2(\alpha+\beta-1).$$

If, then, the discriminant Δ *is not equal to zero, the integral surfaces of* (6) *are projective transformations of the surfaces*

$$X = Y^{\alpha} Z^{\beta},$$

in which the constant exponents α *and* β *are arbitrary except for the restrictions implied by the inequalities* (17) *and* (18). *The surfaces thus excluded either are ruled or have asymptotic lines belonging to linear complexes.*

We shall speak of the surfaces obtained so far as surfaces of the first type. We shall say that a plane anharmonic curve is of the first type, if the one-parameter group of collineations, which leaves it invariant, has an invariant triangle. It is evident that *all plane sections of a surface of the first type are plane anharmonic curves of the first type.*

§ 3. *Discussion of the case when* $c_0 = 0$, *and when two or more roots of the characteristic equations coincide.*

Let α_3 and α_4 be equal and let the other roots of (8) be distinct and different from α_3. Of course, the discriminant Δ must be equal to zero, so that

$$c_1^2 c_2^2 + 4(c_1^3 + c_2^3) + 18 c_1 c_2 - 27 = 0. \tag{19}$$

The three functions

$$y_k = e^{\alpha_k u + \beta_k v} \qquad (k = 1, 2, 3), \tag{20}$$

will still be linearly independent. In accordance with well-known principles, we may choose as a fourth independent solution the function

$$\left[\frac{\partial e^{\alpha u + \beta v}}{\partial \alpha}\right]_{\alpha = \alpha_3}.$$

Since we have

$$\beta = -\tfrac{1}{2}(c_1 + \alpha^2),$$

we find in this way

$$y_4 = (u - \alpha_3 v)\, e^{\alpha_3 u + \beta_3 v}. \tag{21}$$

Introducing non-homogeneous coördinates, we find

$$X = \frac{y_1}{y_3} = e^{(\alpha_1 - \alpha_3)u + (\beta_1 - \beta_3)v}, \qquad Y = \frac{y_2}{y_3} = e^{(\alpha_2 - \alpha_3)u + (\beta_2 - \beta_3)v},$$
$$Z = \frac{y_4}{y_3} = u - \alpha_3 v. \tag{22}$$

In order to be able to eliminate u and v, it is convenient to express α_1, α_2, β_1, β_2 as functions of α_3. We find from equations (8) that

$$\alpha_1 + \alpha_2 = -2\alpha_3, \qquad \alpha_1 \alpha_2 = 2c_1 + 3\alpha_3^2, \tag{23}$$

so that α_1 and α_2 are the two roots of the quadratic

$$\alpha^2 + 2\alpha_3\,\alpha + 2c_1 + 3\alpha_3^2 = 0.$$

But we find further from (8) and (23) that

$$2\alpha_3\,(2c_1 + 3\alpha_3^2) - 2\alpha_3^3 = -8, \qquad (2c_1 + 3\alpha_3^2)\,\alpha_3^2 = c_1^2 + 4c_2,$$

whence

$$\alpha_3^3 + c_1\,\alpha_3 + 2 = 0, \quad 3\alpha_3^4 + 2c_1\,\alpha_3^2 - (c_1^2 + 4c_2) = 0.$$

By means of the first of these equations, the quadratic whose roots are α_1 and α_2 becomes

$$\alpha^2 + 2\alpha_3\,\alpha + \alpha_3^2 - \frac{4}{\alpha_3} = 0,$$

so that we may write

$$\alpha_1 = -\alpha_3 + \frac{2}{\sqrt{\alpha_3}}, \qquad \alpha_2 = -\alpha_3 - \frac{2}{\sqrt{\alpha_3}},$$

where α_3 cannot be equal to zero, since (8) clearly never has zero as a double root. We find

$$\beta_1 = -\tfrac{1}{2}\,(c_1 + \alpha_1^2) = \tfrac{1}{2}\,(\alpha_3^2 - \alpha_1^2) + \frac{1}{\alpha_3},$$

$$\beta_2 = -\tfrac{1}{2}\,(c_1 + \alpha_2^2) = \tfrac{1}{2}\,(\alpha_3^2 - \alpha_2^2) + \frac{1}{\alpha_3},$$

$$\beta_3 = -\tfrac{1}{2}\,(c_1 + \alpha_3^2) = \frac{1}{\alpha_3},$$

so that

$$\alpha_1 - \alpha_3 = 2\left(-\alpha_3 + \frac{1}{\sqrt{\alpha_3}}\right), \qquad \beta_1 - \beta_3 = 2\left(-\frac{1}{\alpha_3} + \sqrt{\alpha_3}\right),$$

(24

$$\alpha_2 - \alpha_3 = 2\left(-\alpha_3 - \frac{1}{\sqrt{\alpha_3}}\right), \qquad \beta_2 - \beta_3 = 2\left(-\frac{1}{\alpha_3} - \sqrt{\alpha_3}\right),$$

and

$$(\alpha_1 - \alpha_3)\,(\beta_2 - \beta_3) - (\alpha_2 - \alpha_3)\,(\beta_1 - \beta_3) = 8\alpha_3^{-\frac{1}{2}}\,(\alpha_3^3 - 1).$$

This latter expression is different from zero. For, if α_3 were equal to a cube-root of unity, either $\alpha_1 - \alpha_3$ or $\alpha_2 - \alpha_3$ would vanish, i. e., α_3 would be a triple root of the characteristic equation, contrary to our assumption. We find, therefore, from (22)

$$4\alpha_3^{-3/2}(\alpha_3^3 - 1)\,u = -\left(\frac{1}{\alpha_3} + \sqrt{\alpha_3}\right)\log X - \left(-\frac{1}{\alpha_3} + \sqrt{\alpha_3}\right)\log Y,$$

$$4\alpha_3^{-3/2}(\alpha_3^3 - 1)\,v = \left(\alpha_3 + \frac{1}{\sqrt{\alpha_3}}\right)\log X + \left(-\alpha_3 + \frac{1}{\sqrt{\alpha_3}}\right)\log Y,$$

and consequently

$$4\alpha_3^{-3/2}(\alpha_3^3 - 1)\, Z = -\frac{(\alpha_3^{3/2}+1)^2}{\alpha_3}\log X + \frac{(\alpha_3^{3/2}-1)^2}{\alpha_3}\log Y,$$

or

$$\frac{4}{\sqrt{\alpha_3}} Z = -\frac{\alpha_3^{3/2}+1}{\alpha_3^{3/2}-1}\log X + \frac{\alpha_3^{3/2}-1}{\alpha_3^{3/2}+1}\log Y,$$

a surface which is projectively equivalent to

$$Z = \log \frac{Y^k}{X^{1/k}}, \tag{25}$$

where k is any finite non-vanishing constant. *We shall speak of* (25) *as a surface of the second type.* Clearly, *only the plane sections* $Z = const.$ *are anharmonic curves of the first type. All other plane sections are anharmonic curves of the second type.*

Consider now the case when the characteristic equation has three equal roots, say

$$\alpha_2 = \alpha_3 = \alpha_4 = \alpha.$$

We shall then have, according to (8),

$$\alpha_1 = -3\alpha, \quad c_1 = -3\alpha^2, \quad \alpha^3 = 1, \quad c_2 = -3\alpha.$$

Our system of differential equations (6) assumes the form

$$y_{uu} + 2y_v - 3\alpha^2 y = 0,$$

$$y_{vv} + 2y_u - 3\alpha y = 0,$$

where α is a cube-root of unity.

By a transformation of the form

$$\bar{u} = \lambda u, \qquad \bar{v} = \mu v,$$

λ and μ being constants, this system becomes

$$\frac{\partial^2 y}{\partial \bar{u}^2} + 2\frac{\mu}{\lambda^2}\frac{\partial y}{\partial \bar{v}} - 3\frac{\alpha^2}{\lambda^2} y = 0,$$

$$\frac{\partial^2 y}{\partial \bar{v}^2} + 2\frac{\lambda}{\mu^2}\frac{\partial y}{\partial \bar{u}} - 3\frac{\alpha}{\mu^2} y = 0,$$

and this is of the same form as the original system if

$$\lambda^2 = \mu, \qquad \mu^2 = \lambda,$$

i. e., if

$$\lambda = 1, \omega, \omega^2; \qquad \mu = 1, \omega^2, \omega,$$

where ω is an imaginary cube root of unity.

Thus, if $\alpha = \omega$, the above transformation for $\lambda = \omega$ will give rise to a new system of the same form for which $\alpha = 1$. Similarly if $\alpha = \omega^2$. We may assume therefore, without any loss of generality, that $\alpha = 1$, so that we have to deal only with the system

$$y_{uu} + 2y_v - 3y = 0, \qquad y_{vv} + 2y_u - 3y = 0. \tag{26}$$

Equation (8) reduces to

$$(\alpha - 1)^3 (\alpha + 3) = 0,$$

so that

$$\alpha_1 = -3, \quad \alpha_2 = \alpha_3 = \alpha_4 = 1,$$

$$\beta_1 = -3, \quad \beta_2 = \beta_3 = \beta_4 = 1.$$

We find the following solutions

$$y_1 = e^{-3(u+v)}, \quad y_2 = e^{u+v}, \quad y_3 = (u - v)e^{u+v}, \tag{27}$$

$$y_4 = [2(u - v)^2 - (u + v)]\, e^{u+v}.$$

Therefore

$$X = \frac{y_3}{y_2} = u - v. \qquad Y = \frac{y_4}{y_2} = 2(u - v)^2 - (u + v), \tag{28}$$

$$Z = \frac{y_1}{y_2} = e^{-4(u+v)},$$

whence

$$Z = e^{4(Y - 2X^2)}, \tag{29}$$

the equation of our surfaces of the third type.

If equation (8) had four equal roots, all of these roots would be equal to zero, since their sum must vanish. But this is clearly impossible, since the sum of their products taken three at a time must be equal to -8. If (8) had two pairs of equal roots we should again encounter the same contradiction. Therefore these two cases are impossible, and *the integral surfaces of system* (6) *must always belong to one of the three types which we have determined.*

§ 4. *General properties of the surfaces of the problem.*

The curves $u = \text{const.}$ and $v = \text{const.}$ are the two families of asymptotic lines of our surface S. The tangents of the curves of the second family, constructed at all of the points of a curve $u = \text{const.}$ of the first family, generate a ruled surface R_1, the osculating ruled surface of the first kind.* If we write

$$z = y_u,$$

* *Second Memoir*, p. 81.

the differential equations of this ruled surface will be

$$y_{vv} + p_{11}\, y_v + p_{12}\, z_v + q_{11}\, y + q_{12}\, z = 0,$$
$$z_{vv} + p_{21}\, y_v + p_{22}\, z_v + q_{21}\, y + q_{22}\, z = 0,$$

where*

$$p_{11} = p_{12} = 0, \qquad q_{11} = c_2, \qquad q_{12} = 2,$$
$$p_{21} = -4, \qquad p_{22} = 0, \qquad q_{21} = -2c_1, \qquad q_{22} = c_2,$$

so that†

$$u_{11} - u_{22} = 0, \qquad u_{12} = -8, \qquad u_{21} = 8c_1,$$
$$v_{11} - v_{22} = -64, \qquad v_{12} = 0, \qquad v_{21} = 0,$$
$$w_{11} - w_{22} = 0, \qquad w_{12} = 0, \qquad w_{21} = 256,$$
$$\theta = (u_{11} - u_{22})^2 + 4u_{12}\, u_{21} = -2^8 c_1, \qquad \Delta = -2^{17}.$$

Since $v_{11} - u_{22} = 0$, *the curves C_y and C_z on R_1 are harmonic conjugates of each other with respect to the two branches of the flecnode curve, which two branches coincide if and only if $c_1 = 0$.*

In general the two branches of the flecnode curve of R_1 are found‡ by factoring the covariant C. In our case, except for a numerical factor, this covariant is equal to $c_1 y^2 + z^2$, so that we find the following parametric expressions for the coördinates of the points of the two branches of the flecnode curve of R_1;

$$x_k = z_k \pm \sqrt{-c_1}\, y_k \qquad (k = 1, 2, 3, 4).$$

For a surface of the first type this becomes

$$x_k = (\alpha_k \pm \sqrt{-c_1})\, e^{\alpha_k u + \beta_k v} = (\alpha_k \pm \sqrt{-c_1})\, y_k.$$

Thus; *each of the branches of the flecnode curve of R_1 is a projective transformation of the corresponding asymptotic curve of S; their locus is a surface of two sheets, each projectively equivalent to S itself; except when one of the roots $\alpha_1, \cdots \alpha_4$ of the characteristic equation satisfy the condition $\alpha^2 + c_1 = 0$, in which case one of the branches of the flecnode curve becomes a plane curve, and the corresponding sheet of the surface which is their locus degenerates into a plane.*

Since our surface S is not a ruled surface, and since no plane curve which is not a straight line can be an asymptotic line of a surface, we see that *a branch of a flecnode curve on one of the osculating ruled surfaces R_1 can never become a plane curve except under the condition just stated.* This may be verified analytically by calculating the Wronskian of $x_1, \cdots x_4$ with respect to v.

* *Loc. cit.*, p. 81.

† *Proj. Diff. Geom.*, pp. 96, 99, 101, 102, 104.

‡ *Proj. Diff. Geom.*, p. 150.

If one of the quantities $\alpha_1, \cdots \alpha_4$ satisfies the condition

$$\alpha_i^2 + c_1 = 0,$$

β_i will vanish, and according to (9) this implies $c_2^2 + 4c_1 = 0$. Thus, *one of the branches of the flecnode curve on the osculating ruled surfaces R_1 of S will be a plane curve, if and only if*

$(30)_1$ $$c_2^2 + 4c_1 = 0.$$

The corresponding condition for the osculating ruled surfaces R_2 of the second kind is

$(30)_1$ $$c_1^2 + 4c_2 = 0.$$

If both conditions are satisfied we must have for c_1, c_2 one of the four sets of values

$$c_1 = 0, \quad -4, \quad -4\omega, \quad -4\omega^2,$$

$$c_2 = 0, \quad -4, \quad -4\omega^2, \quad -4\omega,$$

where ω is a complex cube root of unity. As in the preceding paragraph these cases may be reduced to two, viz.:

$$c_1 = c_2 = 0 \quad \text{or} \quad c_1 = c_2 = -4.$$

Those surfaces of the first type all of whose osculating ruled surfaces, of both kinds, have a flecnode curve with a plane branch are integral surfaces of a system of form (6) *in which either $c_1 = c_2 = 0$ or $c_1 = c_2 = -4$. The flecnode curves of the osculating ruled surfaces of the surfaces belonging to the first of these two classes, $(c_1 = c_2 = 0)$, coincide entirely with the plane branch. For these surfaces the two branches of the flecnode curve coincide. This is not the case for the surfaces of the second class $(c_1 = c_2 = -4)$.*

These two cases merit closer attention. We have, in the first case, $c_1 = c_2 = 0$ and therefore

$$y_1 = e^{-2(u+v)}, \quad y_2 = e^{-2(\omega u + \omega^2 v)}, \quad y_3 = e^{-2(\omega^2 u + \omega v)}, \quad y_4 = 1,$$

where

$$\omega = -\tfrac{1}{2} + \tfrac{1}{2} i\sqrt{3}, \qquad \omega^2 = -\tfrac{1}{2} - \tfrac{1}{2} i\sqrt{3},$$

so that

$$y_2 = e^{u+v+i\sqrt{3}(-u+v)}, \qquad y_3 = e^{u+v+i\sqrt{3}(u-v)}.$$

Let us introduce non-homogeneous coördinates by means of the projective transformation

$$\xi = \frac{y_2 + y_3}{2y_1}, \qquad \eta = \frac{y_3 - y_2}{2iy_1}, \qquad \zeta = \frac{y_4}{y_1}.$$

We find

$$\xi = e^{3(u+v)} \cos \sqrt{3}\,(u-v),$$

$$\eta = e^{3(u+v)} \sin \sqrt{3}\,(u-v), \tag{31}$$

$$\zeta = e^{2(u+v)},$$

whence we obtain

$$\zeta^3 = \xi^2 + \eta^2 \tag{32}$$

as the equation of our surface. If we interpret ξ, η, ζ as Cartesian coördinates, as we may, *this is a surface of revolution generated by rotating the semi-cubical parabola*

$$\xi^2 = \zeta^3$$

about the ζ axis.

The asymptotic lines of the surface are obtained by equating either u or v in (31) to a constant. *The projections of these curves on the $\xi\eta$ plane are logarithmic spirals which intersect the radii vectores at a constant angle of* 60°. *They are, therefore, curves all of whose points are coincidence points.**

We know already that the flecnode curves of the osculating ruled surfaces have coincident branches, and are plane curves. It is easy to show that these plane curves coincide with the above two families of logarithmic spirals. More specifically we may express our result as follows.

Every asymptotic line A of the surface

$$\zeta^3 = \xi^2 + \eta^2 \tag{S}$$

has as its projection in the $\xi\eta$ plane a logarithmic spiral intersecting all radii vectores at an angle of 60°. *The flecnode curve of the ruled surface which osculates S along A is a logarithmic spiral obtained from the projection of A by a rotation of* 120° *around the origin.*

A somewhat similar situation exists for surfaces of the second class, for which $c_1 = c_2 = -4$. We find, in this case, the fundamental solutions

$$y_1 = e^{2v}, \quad y_2 = e^{2u}, \quad y_3 = e^{(-1+\sqrt{5})(u+v)}, \quad y_4 = e^{-(1+\sqrt{5})(u+v)}. \tag{33}$$

If we put

$$X = \frac{y_1}{y_4}, \qquad Y = \frac{y_2}{y_4}, \qquad Z = \frac{y_3}{y_4},$$

we find

$$Z = (XY)^{5-2\sqrt{5}} \tag{34}$$

as the equation of the surface, with

$$X = e^{(1+\sqrt{5})u+(3+\sqrt{5})v}, \qquad Y = e^{(3+\sqrt{5})u+(1+\sqrt{5})v}, \qquad Z = e^{2\sqrt{5}(u+v)} \tag{35}$$

as its parametric representation.

* *Proj. Diff. Geom.*, p. 70.

Consider an asymptotic line $v = \text{const.}$ of this surface and the osculating ruled surface R_1 which corresponds to it. The homogeneous coördinates x_k' and x_k'' of the points which describe the two branches of the flecnode curve of R_1 will be

$$x_1' = 2y_1', \quad x_2' = 4y_2, \quad x_3' = (1+\sqrt{5})\,y_3, \quad x_4' = -(-1+\sqrt{5})\,y_4,$$

$$x_1'' = -2y_1, \quad x_2'' = 0, \quad x_3'' = (-3+\sqrt{5})y_3, \quad x_4'' = -(3+\sqrt{5})\,y_4.$$

Thus *the second branch of the flecnode curve is indeed a plane curve in the XZ plane, the non-homogeneous coördinates of whose points are given by*

$$X'' = \frac{2}{3+\sqrt{5}}X, \qquad Y'' = 0, \qquad Z'' = \frac{3-\sqrt{5}}{3+\sqrt{5}}Z.$$

These curves clearly coincide with the projections on the XZ plane of the asymptotic curves $v = \text{const.}$ of the original surface. An analogous relation exists between the plane branches of the flecnode curves of the osculating ruled surfaces R_2, and the projections on the YZ plane of the asymptotic curves $u = \text{const.}$

For surfaces S of the second type we have, of course

$$c_1^2 c_2^2 + 4(c_1^3 + c_2^3) + 18c_1c_2 - 27 = 0. \tag{19}$$

The locus of the flecnode curves of the osculating ruled surfaces R_1 is again a surface of two sheets, each of which is projective to S itself, unless one of the three distinct roots $\alpha_1, \alpha_2, \alpha_3$ of the characteristic equation satisfies the condition

$$\alpha^2 + c_1 = 0,$$

This, again, can happen only if

$$c_2^2 + 4c_1 = 0. \tag{30$_1$}$$

The conditions (19) and (30)$_1$ yield

$$c_1 = -\frac{9}{\sqrt[3]{16}}, \qquad c_2 = -3\sqrt[3]{2},$$

the other possibilities, involving imaginary cube roots of unity as factors being disposed of as in the preceding paragraph. The double root α_3 of the characteristic equation will be

$$\alpha_3 = \frac{2(c_1^2 + 3c_2)}{c_1c_2 - 9} = \frac{9}{25\sqrt[3]{4}},$$

whence

$$\alpha_2^{3/2} = \pm\frac{27}{250}, \qquad k = \frac{\alpha^{3/2}-1}{\alpha^{3/2}+1} = -\frac{223}{277} \quad \text{or} \quad -\frac{277}{223}.$$

Substituting these values in (25) we find the following result. *Any surface of the second type whose osculating ruled surfaces have flecnode curves one of whose branches is a plane curve is a projective transformation of the surface*

$$Z = \log \frac{X^{\overline{111}}}{Y^{\overline{111}}}. \tag{36}$$

The flecnode curve of an osculating ruled surface of a surface of the third type never has a plane branch, and both of its branches are projective transformations of the corresponding asymptotic curve of the surface itself.

If we use the four points

$$y, \quad z = y_u, \quad \rho = y_v, \quad \sigma = y_{uv}$$

as tetrahedron of reference, the equations of the osculating linear complexes of R_1 and R_2 reduce to

$$\begin{aligned} \omega_{13} - 2\omega_{34} + c_1\,\omega_{42} &= 0, \\ \omega_{12} - 2\omega_{42} - c_2\,\omega_{34} &= 0 \end{aligned} \tag{37}$$

respectively, and these are special if and only if c_1 or c_2 is equal to zero. The two complexes are, of course, in involution.* The osculating linear complexes of the asymptotic curves Γ' (v = const.) and Γ'' (u = const.) are respectively

$$-\omega_{14} + \omega_{23} = 0, \qquad \omega_{14} + \omega_{23} = 0. \tag{38}$$

As distinguished from the general theory, *any two of these four linear complexes are in involution.*

The osculating linear complexes of Γ' and Γ'' have in common a linear congruence. According to the general theory† one of the directrices d of this congruence (that of the first kind) lies in the tangent plane of the surface point considered, while the other d' (that of the second kind), passes through this point. Their equations are

$$x_1 = x_4 = 0 \quad \text{and} \quad x_2 = x_3 = 0$$

respectively, so that d coincides with $P_z P_\rho$ and d' with $P_y P_\sigma$.

Since we have a line d and a line d' associated with every point of the surface, we thus obtain two congruences, the *directrix congruences of the first and second kind.* The curves on S which correspond to the developables of these congruences are called the *directrix curves* of S, and are the same for both congruences. Their differential equation, in our case, reduces to‡

$$c_1\,du^2 - c_2\,dv^2 = 0.$$

* *Second Memoir*, p. 90.

† *Loc. cit.*, p. 95.

‡ *Loc. cit.*, pp. 116 and 119.

Thus, *the directrix curves of S form a conjugate system and are obtained by putting* $\bar{u} = const.$ *and* $\bar{v} = const.$, *where*

$$\bar{u} = \sqrt{c_1}\,u + \sqrt{c_2}\,v, \qquad \bar{v} = \sqrt{c_1}\,u - \sqrt{c_2}\,v, \tag{39}$$

if c_1 *and* c_2 *are both different from zero. If one only of these quantities vanishes, the directrix curves of both families coincide with one of the two families of asymptotic lines. If* c_1 *and* c_2 *both vanish, the directrix curves are indeterminate.*

If c_1 and c_2 are both different from zero, and if the surface S is of the first type, its parametric equations referred to its directrix curves will be

$$y_k = e^{\frac{1}{2}\left(\frac{a_k}{\sqrt{c_1}} + \frac{\beta_k}{\sqrt{c_2}}\right)\bar{u} + \frac{1}{2}\left(\frac{a_k}{\sqrt{c_1}} - \frac{\beta_k}{\sqrt{c_2}}\right)\bar{v}} \qquad (k = 1, 2, 3, 4). \tag{40}$$

Consequently *the directrix curves of S are also anharmonic.*

If the point P moves along the directrix curve $\bar{u} = \text{const.}$, the line d will describe a developable of the directrix congruence of the first kind. It will meet the cuspidal edge of this developable in the point

$$\lambda' z + \mu' \rho,$$

where*

$$\lambda' : \mu' = \sqrt{c_2} : \sqrt{c_1}.$$

Consequently, *if in the expressions*

$$\eta_k = \sqrt{c_2}\, z_k + \sqrt{c_1}\, \rho_k = (\alpha_k \sqrt{c_2} + \beta_k \sqrt{c_1})\, y_k \quad (k = 1, 2, 3, 4), \tag{41}$$

we put $\bar{u} = const.$, *we obtain the cuspidal edge of the developable described by d when P moves along a curve* $\bar{u} = const.$ *of a surface of the first type. If we allow* $\bar{u}$ *to vary, we obtain the locus of these cuspidal edges, i. e., one of the sheets of the focal surface of the directrix congruence of the first kind. The other sheet will be given by*

$$\zeta_k = (\alpha_k \sqrt{c_2} - \beta_k \sqrt{c_1})\, y_k \qquad (k = 1, 2, 3, 4). \tag{42}$$

In general, both of these sheets are surfaces which are projectively equivalent to S itself. The only exceptions to this statement will arise when at least one of the number pairs a_k, β_k satisfies the equation

$$c_2\, \alpha_k^2 - c_1\, \beta_k^2 = 0.$$

But α_k and β_k are solutions of (7). Therefore in this case we should also have

$$c_1\, \alpha_k - c_2\, \beta_k = 0.$$

We are assuming that c_1 and c_2 are both different from zero. The last two equations then imply

$$a_k^3 = \beta_k^3, \qquad c_1^3 = c_2^3,$$

* *Loc. cit.*, p. 119.

and, as we have seen in several instances, we may substitute for this later condition the simpler one

$$c_1 = c_2,$$

without loss of generality.

On the other hand, if $c_1 = c_2$, we find from (7) that

$$a^2 - \beta^2 + 2(\beta - \alpha) = 0$$

or

$$(\alpha - \beta)(\alpha + \beta - 2) = 0.$$

Therefore two of the roots α_1, α_2, of the characteristic equation (8) will satisfy the quadratic

$$\alpha^2 + 2\alpha + c_1 = 0$$

and we shall have $\beta_1 = \alpha_1$, $\beta_2 = \alpha_2$. The other two roots α_3, α_4 will be solutions of

$$\alpha^2 + 2(2 - \alpha) + c_1 = 0$$

and we shall have

$$\beta_3 = 2 - \alpha_3, \qquad \beta_4 = 2 - \alpha_4.$$

The assumption $\beta_3 = \alpha_3$ would imply $\alpha_3 = 1$, and therefore $c_1 = c_2 = -3$. But in this case equation (8) would have three equal roots and the surface S would not be of the first type.

Therefore, *each of the two sheets of the focal surface of the directrix congruence of the first kind is a projective transformation of the surface S itself if $c_1^3 \neq c_2^3$. If $c_1^3 = c_2^3 \neq 0$, one of the sheets of the focal surface degenerates into a straight line. If $c_1 = c_2 = 0$, both sheets degenerate into coplanar straight lines and the directrix curves of S are indeterminate.**

The general properties of the congruence may be studied by means of a system of partial differential equations satisfied by the four pairs of functions (η_k, ζ_k).† This system consists of two partial differential equations of the first, and two of the second order. In order to obtain these equations, let us form $\partial\eta_k / \partial\bar{v}$ and $\partial\zeta_k / \partial\bar{u}$. We find

$$\frac{\partial \eta_k}{\partial \bar{v}} = \tfrac{1}{2}(\alpha_k \sqrt{c_2} + \beta_k \sqrt{c_1}) \left(\frac{\alpha_k}{\sqrt{c_1}} - \frac{\beta_k}{\sqrt{c_2}} \right) y_k = \frac{\alpha_k^2 c_2 - \beta_k^2 c_1}{2\sqrt{c_1 c_2}} y_k,$$

or, making use of (7),

$$\frac{\partial \eta_k}{\partial \bar{v}} = \frac{c_1 \alpha_k - c_2 \beta_k}{\sqrt{c_1 c_2}} y_k.$$

* This theorem has here actually been proved only for surfaces of the first type. It may be shown to be true in the other cases also.

† E. J. WILCZYNSKI, *Sur la théorie générale des congruences*, Mémoires publies par la Classe des Sciences de l'Académie royale de Belgique, collection en 4°, Deuxieme série, tome III (1911).

We have further, from (41) and (42),

$$2\alpha_k \sqrt{c_2}\, y_k = \eta_k + \zeta_k, \qquad 2\beta_k \sqrt{c_1}\, y_k = \eta_k - \zeta_k,$$

and therefore

$$\frac{\partial \eta_k}{\partial \bar{v}} = -\mu\eta_k + \lambda\zeta_k \qquad (k = 1, 2, 3, 4),$$

where

$$\lambda = \frac{c_2^3 + c_1^3}{2c_1 c_2}, \qquad \mu = \frac{c_2^3 - c_1^3}{2c_1 c_2}. \tag{43}$$

In precisely the same way we find the second equation of the following system:

$$\frac{\partial \eta_k}{\partial \bar{v}} = \frac{\partial \zeta_k}{\partial \bar{u}} = -\mu\eta_k + \lambda\zeta_k. \tag{44}$$

If we put

$$\eta = e^{-\mu\bar{v}}\mathfrak{y}, \qquad \zeta = e^{\lambda\bar{u}}\mathfrak{z}, \tag{45}$$

we find that $\mathfrak{y}_k$ and $\mathfrak{z}_k$ are simultaneous solutions of the equations

$$\frac{\partial \mathfrak{y}}{\partial \bar{v}} = m\mathfrak{z}, \qquad \frac{\partial \mathfrak{z}}{\partial \bar{u}} = n\mathfrak{y}, \tag{46}$$

where

$$m = \lambda e^{\lambda\bar{u}+\mu\bar{v}}, \qquad n = -\mu e^{-\lambda\bar{u}-\mu\bar{v}}. \tag{47}$$

The equations (46) are the required first order equations of our system.

In order to find the second order equations of our system, we observe in the first place that $\eta_1, \cdots \eta_4$, being linear homogeneous functions of $y_1, \cdots y_4$ with constant coefficients, satisfy the same system of partial differential equations as the latter functions, viz.:

$$\eta_{uu} + 2\eta_v + c_1\eta = 0, \qquad \eta_{vv} + 2\eta_u + c_2\eta = 0. \tag{48}$$

Since

$$\bar{u} = \sqrt{c_1}\, u + \sqrt{c_2}\, v, \qquad \bar{v} = \sqrt{c_1}\, u - \sqrt{c_2}\, v,$$

we find

$$\eta_u = \sqrt{c_1}\,(\eta_{\bar{u}} + \eta_{\bar{v}}), \qquad \eta_v = \sqrt{c_2}\,(\eta_{\bar{u}} - \eta_{\bar{v}}),$$

$$\eta_{uu} = c_1(\eta_{\bar{u}\bar{u}} + 2\eta_{\bar{u}\bar{v}} + \eta_{\bar{v}\bar{v}}),$$

$$\eta_{uv} = \sqrt{c_1 c_2}\,(\eta_{\bar{u}\bar{u}} - \eta_{\bar{v}\bar{v}}),$$

$$\eta_{vv} = c_2(\eta_{\bar{u}\bar{u}} - 2\eta_{\bar{u}\bar{v}} + \eta_{\bar{v}\bar{v}}).$$

Consequently we obtain for η as function of $\bar{u}$, $\bar{v}$ the following system of partial differential equations

$$\begin{aligned} \eta_{\bar{u}\bar{u}} + \eta_{\bar{v}\bar{v}} + 2\lambda\eta_{\bar{u}} - 2\mu\eta_{\bar{v}} + \eta = 0, \\ \eta_{\bar{u}\bar{v}} + \mu\eta_{\bar{u}} - \lambda\eta_{\bar{v}} = 0. \end{aligned} \tag{49}$$

Of course $\zeta_1, \cdots \zeta_4$ satisfy the same equations. The transformation (45) now yields the equations

$$\begin{gathered}
\mathfrak{y}_{uu} + \mathfrak{y}_{vv} + 2\lambda\mathfrak{y}_u - 4\mu\mathfrak{y}_v + (1 + 3\mu^2)\,\mathfrak{y} = 0, \\
\mathfrak{y}_{uv} - \lambda\mathfrak{y}_v + \lambda\mu\mathfrak{y} = 0, \\
\mathfrak{z}_{uu} + \mathfrak{z}_{vv} + 4\lambda\mathfrak{z}_u - 2\mu\mathfrak{z}_v + (1 + 3\lambda^2)\,\mathfrak{z} = 0, \\
\mathfrak{z}_{uv} + \mu\mathfrak{z}_u + \lambda\mu\mathfrak{z} = 0.
\end{gathered} \tag{50}$$

But from (46) we find

$$\mathfrak{y}_v = m\mathfrak{z}, \qquad \mathfrak{z}_u = n\mathfrak{y},$$

$$\mathfrak{y}_{vv} = m_v\,\mathfrak{z} + m\mathfrak{z}_v, \qquad \mathfrak{z}_{uu} = n_u\,\mathfrak{y} + n\mathfrak{y}_u.$$

so that, for the first and third equation of system (50), we may substitute the following:

$$\begin{aligned}
\mathfrak{y}_{uu} &= -(1 + 3\mu^2)\,\mathfrak{y} + 3\mu m\mathfrak{z} - 2\lambda\mathfrak{y}_u - m\mathfrak{z}_v, \\
\mathfrak{z}_{vv} &= -3\lambda n\mathfrak{y} - (1 + 3\lambda^2)\,\mathfrak{z} - n\mathfrak{y}_u + 2\mu\mathfrak{z}_v.
\end{aligned} \tag{51}$$

Equations (46) *and* (51) *together constitute the system of partial differential equations characteristic of the directrix congruence of the first kind.* We have actually proved this for the case of a surface of the first type, but the statement remains true for surfaces of the second and third type as well.

The partial differential equations of the congruence show that *it is a W-congruence.** This also follows from the fact that the point correspondence between the two sheets of its focal surface, which is established by the lines of the congruence, is a projective transformation.

The line d, of the congruence, which belongs to the values u, v of the parameters, is obtained by joining the points P_z and P_ρ, where

$$z = y_u, \qquad \rho = y_v.$$

If the surface S is of the first type,

$$\alpha_1 y_1, \quad \alpha_2 y_2, \quad \alpha_3 y_3, \quad \alpha_4 y_4$$

and

$$\beta_1 y_1, \quad \beta_2 y_2, \quad \beta_3 y_3, \quad \beta_4 y_4$$

will, therefore, be the coördinates of two points of d. The Plückerian line coördinates of d will therefore be equal to

$$\omega_{ik} = l_{ik}\, y_i\, y_k \qquad (i, k = 1, 2, 3, 4),$$

where

$$l_{ik} = \alpha_i \beta_k - \alpha_k \beta_i.$$

* *Loc. cit.*, p. 56.

Consequently the coördinates of d will satisfy the quadratic equations

$$l_{13}\, l_{42}\, \omega_{12}\, \omega_{34} - l_{12}\, l_{34}\, \omega_{13}\, \omega_{42} = 0,$$

$$l_{13}\, l_{42}\, \omega_{14}\, \omega_{23} - l_{14}\, l_{23}\, \omega_{13}\, \omega_{42} = 0,$$

$$l_{14}\, l_{23}\, \omega_{12}\, \omega_{34} - l_{12}\, l_{34}\, \omega_{14}\, \omega_{23} = 0,$$

which, on account of the identical relations

$$\omega_{12}\, \omega_{34} + \omega_{13}\, \omega_{42} + \omega_{14}\, \omega_{23} = 0,$$

$$l_{12}\, l_{34} + l_{13}\, l_{42} + l_{14}\, l_{23} = 0,$$

are equivalent to a single one of them. Therefore, *the lines of the directrix congruence of the first kind, of a surface of the first type, belong to a tetrahedral complex whose fundamental tetrahedron coincides with the tetrahedron left invariant by the collineation group of the surface.* The invariant of this tetrahedral complex may be written in the form

$$\frac{l_{12}\, l_{34}}{l_{13}\, l_{42}}.$$

This theorem is modified for surfaces of the second type. *The group of collineations which leave invariant a surface of the second type has no invariant tetrahedron. The quadratic complex to which the directrix congruence of the first kind belongs in this case has the equation*

$$l_{13}\, l_{23}\, \omega_{12}\, \omega_{34} + l_{12}\, (\alpha_3^2 + \beta_3)\, \omega_{13}\, \omega_{23} = 0.$$

It consists of the lines which meet corresponding pairs of lines of two projective pencils, the plane of one (and only one) of these pencils passing through the vertex of the other.

The directrix congruence of the first kind, of a surface of the third type, does not belong to any non-degenerate quadratic complex. Its lines belong, instead, to a unique linear complex.

Let us now consider the directrix congruence of the second kind. If P describes one of the directrix curves on S, the directrix d' will describe one of the developables of this congruence, and will meet the cuspidal edge of this developable in the point

$$\lambda y + 2\mu\sigma,$$

where*

$$\frac{\lambda}{\mu} = -8 + 2c_2 \frac{dv}{du}.$$

Thus, if P describes the curve $\bar{u} = \text{const.}$, d' generates a developable whose

* *Second Memoir*, p. 117.

cuspidal edge is given by

$$2\eta_k' = (-8 - 2\sqrt{c_1 c_2})\, y_k + 2\sigma_k = 2\,(\alpha_k \beta_k - 4 - \sqrt{c_1 c_2})\, y_k. \tag{52}$$

Similarly, the cuspidal edge of the developable described by d', when P moves along the curve $\bar{v} = \text{const.}$, is given by

$$\zeta_k' = (\alpha_k \beta_k - 4 + \sqrt{c_1 c_2})\, y_k \qquad (k = 1, 2, 3, 4). \tag{53}$$

The surfaces $S_{\eta'}$ and $S_{\zeta'}$ constitute the focal surface of the congruence. They are projective transformations of S itself and of each other, unless at least one pair of the quantities α_k, β_k satisfies the equation

$$(\alpha\beta - 4)^2 = c_1 c_2,$$

which according to (7) is equivalent to

$$4\alpha\beta - 2\,(c_1 \alpha + c_2 \beta) - 16 = 0.$$

But this equation, together with (7), implies $c_1^3 = c_2^3$. Thus *the reduction of one of the sheets of the focal surface to a straight line takes place for both directrix congruences simultaneously if* $c_1^3 = c_2^3$, *and in no other case.*

We find

$$\frac{\partial \eta_k'}{\partial \bar{v}} = \tfrac{1}{2}\,(\alpha_k \beta_k - 4 - \sqrt{c_1 c_2})\left(\frac{\alpha_k}{\sqrt{c_1}} - \frac{\beta_k}{\sqrt{c_2}}\right) y_k = \frac{c_2^{\frac{3}{2}} - c_1^{\frac{3}{2}}}{\sqrt{c_1 c_2}}\, y_k,$$

$$\frac{\partial \zeta_k'}{\partial \bar{u}} = \tfrac{1}{2}\,(\alpha_k \beta_k - 4 + \sqrt{c_1 c_2})\left(\frac{\alpha_k}{\sqrt{c_1}} + \frac{\beta_k}{\sqrt{c_2}}\right) y_k = \frac{c_2^{\frac{3}{2}} + c_1^{\frac{3}{2}}}{\sqrt{c_1 c_2}}\, y_k,$$

and

$$\zeta_k' - \eta_k' = 2\sqrt{c_1 c_2}\, y_k,$$

so that we have

$$\frac{\partial \eta_k'}{\partial \bar{v}} = \mu\,(\zeta_k' - \eta_k'), \qquad \frac{\partial \zeta_k'}{\partial \bar{u}} = \lambda\,(\zeta_k' - \eta_k'),$$

making use of the notations (43). Let us assume λ and μ both different from zero, and put

$$\eta_k' = \mu e^{-\mu \bar{v}}\, \mathfrak{y}_k', \qquad \zeta_k' = \lambda e^{\lambda \bar{u}}\, \mathfrak{z}_k', \tag{54}$$

and we shall find

$$\frac{\partial \mathfrak{y}_k'}{\partial \bar{v}} = m\mathfrak{z}_k', \qquad \frac{\partial \mathfrak{z}_k'}{\partial \bar{u}} = n\mathfrak{y}_k', \tag{55}$$

where m and n have the same values as in (47).

On account of (52) and (53), η_k' and ζ_k' satisfy the same system of partial differential equations (6) of which $y_1, \cdots y_4$ are solutions. If we introduce $\bar{u}$ and $\bar{v}$ as independent variables, we find that both $\eta_1', \cdots \eta_4'$ and $\zeta_1', \cdots \zeta_4'$ are solutions of (49). Since the transformations (54) differ from (45) merely by

constant factors, we find that $\mathfrak{y}'_k$ and $\mathfrak{z}'_k$ satisfy equations (50), so that we shall have in particular

$$\mathfrak{y}'_{uu} = -\mathfrak{y}'_{vv} - 2\lambda\mathfrak{y}'_u + 4\mu\mathfrak{y}'_v - (1+3\mu^2)\,\mathfrak{y}',$$

$$\mathfrak{z}'_{vv} = -\mathfrak{z}'_{uu} - 4\lambda\mathfrak{z}'_u + 2\mu\mathfrak{z}'_v - (1+3\lambda^2)\,\mathfrak{z}'.$$

If we make use of (55), we find

$$\begin{aligned} \mathfrak{y}_{uu} &= -(1+3\mu^2)\,\mathfrak{y}' + 3\mu m\mathfrak{z}' - 2\lambda\mathfrak{y}'_u - m\mathfrak{z}'_v, \\ \mathfrak{z}'_{vv} &= -3\lambda n\mathfrak{y}' - (1+3\lambda^2)\,\mathfrak{z}' - n\mathfrak{y}'_u + 2\mu\mathfrak{z}'_v. \end{aligned} \tag{56}$$

Equations (55) and (56) are precisely the same as (46) and (51) except for the letters used for the unknown functions. Consequently *the two directrix congruences are equivalent to each other by means of a projective transformation.* If λ or μ should vanish, the transformation (54) ceases to be available, but the two congruences are nevertheless projectively equivalent. Since, in this case, we may assume $c_1 = c_2 = c$, the discriminant Δ becomes equal to

$$c^4 + 8c^3 + 18c^2 - 27 = (c+3)^3\,(c-1).$$

The surface will be of the first type only if c is neither unity nor minus three.

THE UNIVERSITY OF CHICAGO,
August 27, 1912.

ON THE REPRESENTATION GROUPS OF GIVEN ABSTRACT GROUPS*

BY

G. A. MILLER

§ 1. *Introduction.*

The abstract group G is said to be a *representation group* (Darstellungsgruppe) of the finite abstract group G' whenever G is one of the largest groups involving a quotient group which is simply isomorphic with G' as regards a subgroup M which is contained both in the central and also in the commutator subgroup of G. A number of fundamental theorems relating to representation groups were developed by I. SCHUR in two memoirs published in the Journal für Mathematik.† J. A. DE SEGUIER suggested some simplifications relating to these memoirs.‡ The objects of the present article are to establish a few additional theorems bearing directly on this subject, and to exhibit other points of contact with known developments. It is hoped that the method of approach here adopted will also throw additional light on some of the known results relating to this important system of groups. In particular, we shall not employ the theory of group characters, which was used by I. Schur.

When G' is abelian it results directly from the definition of the representation groups that the group of inner isomorphisms of G is abelian and hence G must be either abelian or metabelian.§ As every metabelian group is known to be the direct product of its Sylow subgroups, it follows that every representation group G of an abelian group is the direct product of the Sylow subgroups of G. Every invariant subgroup whose complementary group is abelian is known to be contained in the commutator subgroup. Hence M is the commutator subgroup of G whenever G is abelian, and as the commutator subgroup of a direct product is the direct product of the commutator subgroups of the factors, G is the direct product of representation groups of the Sylow subgroups of G'. That is, *a representation group of any abelian group is a direct*

* Presented to the Society (Chicago), March 21, 1913.

† Vol. 127 (1904), p. 20, and vol. 132 (1907), p. 85.

‡ Paris Comptes Rendus, vol. 145 (1907), p. 303.

§ A metabelian group is a non-abelian group whose group of inner isomorphisms is abelian. W. BURNSIDE, *Theory of Groups of Finite Order*, second edition, 1911, p. 57, uses this term with a more general meaning.

product of representation groups of the Sylow subgroups of this abelian group. This theorem is therefore a special result from known theorems on metabelian groups and on commutator subgroups.

Since the group of inner isomorphisms of a group cannot be cyclic* and since every quotient group of a cyclic group is cyclic, it results that the representation group of a cyclic group is simply isomorphic with this cyclic group. In other words, when G is cyclic, M reduces to the identity. It is easy to see that *a necessary and sufficient condition that an abelian group is its own representation group is that this abelian group is cyclic,* and that a non-cyclic abelian group has always at least two different representation groups. In particular, the representation groups of the non-cyclic group of order p^2 are the two non-abelian groups of order p^3, since the orders of the commutators of a metabelian group must divide the orders of operators in its group of inner isomorphisms.†

§ 2. *Commutators and commutator subgroups.*

The commutator of the two operators s and t will be written in the form $s^{-1}t^{-1}st = c$. Since

$$(s^{-1}t^{-1}s)^n = (ct^{-1})^n = ct^{-1}ct \cdot t^{-2}ct^2 \cdot \cdots \cdot t^{-(n-1)}ct^{n-1} \cdot t^{-n},$$

there results the elementary theorem that if the nth power of one element of a commutator is commutative with the other element, then the continued product of the n conjugates of their commutator under the former element is the identity. In particular, *If the square of an operator is commutative with another operator, their commutator is transformed into its inverse by the former of these two operators. If a commutator is commutative with one of its two elements its order is a divisor of the index of the lowest power to which this element must be raised to be commutative with the other element.*‡

The wide application of these elementary theorems may be inferred from the fact that every possible non-abelian group of finite order contains at least one commutator, besides the identity, which is commutative with at least one of the elements of the commutator. In the case of a non-abelian group whose commutator subgroup is abelian this theorem is almost self-evident. If such a group is metabelian each of its commutators is commutative with all of the operators of the group. If it is not abelian, the commutators formed by an operator in the commutator subgroup and any operator which is not commutative with it will have the required property. As every non-abelian

*Paris Comptes Rendus, vol. 128 (1889), p. 229.

†W. B. Fite, These Transactions, vol. 3 (1902), p. 334.

‡Bulletin of the American Mathematical Society, vol. 4 (1898), p. 137.

solvable group must involve a non-abelian group whose commutator subgroup is abelian, it results that every solvable group must involve a commutator which satisfies the condition in question.

It remains to prove that every insolvable group has a commutator, besides the identity, which is commutative with at least one of its elements. If there were an insolvable group which did not have this property, there would be such a group of smallest order. Each subgroup of this group would therefore be abelian, otherwise such a subgroup would contain a commutator which would satisfy the condition in question. As every insolvable group is known to contain at least one non-abelian subgroup,* we arrive at a contradiction by assuming that an insolvable group exists which does not contain a commutator besides the identity, which is commutative with at least one of its elements. This completes a proof of the theorem: *Every possible non-abelian group of finite order contains at least one commutator besides the identity which is commutative with at least one of the elements of the commutator.*

If both of the elements of a commutator are commutative with this commutator, the metabelian group generated by these two elements has a cyclic commutator subgroup, and this commutator subgroup is generated by the commutator of these two elements. It is known that if a commutator is commutative with one of its elements, the αth power of this commuatator is the commutator of the αth power of this element and the other element of the given commutator.† Hence it results directly that the group generated by the commutator of these two elements is the commutator subgroup of the group generated by the elements. As an illustration of this elementary theorem, we may observe that if G' is an abelian group having two invariants (m_1, m_2) the representation group of G' has a cyclic commutator subgroup whose order is the highest common factor of the two numbers m_1 and m_2. The order of this G is therefore the product of $m_1 m_2$ and this highest common factor.

The number of the distinct commutators in the representation group of a group G' of order g' is evidently always less than g'^2, since a commutator is not affected if one or both of its elements are multiplied by operators which are commutative with these elements. Hence a commutator is not changed if either one of its elements is replaced by any other element of G which is found in the same co-set as the given element, with respect to M. This is clearly a special case of the elementary theorem that *the number of the distinct commutators of a group is always less than the square of the order of its central quotient group.* Although this theorem can easily be generalized, this special case is sufficient for our object in view in the next section, viz., to establish an infinite system

* These Transactions, vol. 4 (1903), p. 398.

† Bulletin of the American Mathematical Society, vol. 4 (1898), p. 137.

of groups whose commutator subgroups involve operators which are not commutators.

To exhibit some of the applications of the elementary theorems of this section we may consider the representation group of the icosahedral group. As each of the operators of this group is transformed into its inverse by an operator of order 2 contained in the group, it results that every group which has the icosahedral group as a central quotient group must be such that each of its co-sets which corresponds to an operator of orders 3 in the icosahedral group must involve at least one commutator of order 3, while each of those which correspond to operators of order 5 must involve at least one commutator of this order. Since each of the operators of order 2 in the icosahedral group is transformed into itself multiplied by an operator of order 2, it results that the co-sets which correspond to operators of order 2 in the given isomorphisms must involve commutators which are transformed into their inverses, and hence their orders must divide 4. In particular, the representation group of the icosahedral group is such that it involves operators of orders 4, 3, and 5 in the co-sets which correspond to operators of orders 2, 3, and 5 respectively in this icosahedral group. As a special case we have the result that *the representation group of the icosahedral group is of order* 120.* This group of order 120 may be distinguished by the facts that it contains only one operator of order 2 and is isomorphic with the icosahedral group.

§ 3. *Representation groups of non-cyclic abelian groups.*

From the theorems of the preceding section, especially from those of the last three paragraphs, it is very easy to find the order of the representation groups G of any non-cyclic abelian group G'. Suppose that the invariants of G' are $m_1, m_2, \cdots, m_\lambda$, and that each of these invariants is divisible by all those which follow it. The commutator subgroup M of G will clearly contain $\lambda - 1$ invariants which are equal to m_1, $\lambda - 2$ which are equal to $m_2, \cdots$, and one which is equal to $m_{\lambda-1}$. Hence the order of M is given by the formula

$$m_1^{\lambda-1}\, m_2^{\lambda-2} \cdots m_{\lambda-1},$$

and the order of G is

$$m_1^{\lambda}\, m_2^{\lambda-1} \cdots m_{\lambda-1}^{2}\, m_\lambda.\dagger$$

The number of the invariants of M is therefore $\frac{1}{2}\lambda\,(\lambda - 1)$, if we assume that these invariants have been so chosen that they can be arranged linearly so that each of them is a divisor of all those which follow it. In what follows we shall always assume that the invariants of an abelian group have been selected in this manner, unless the contrary is explicitly stated.

*Quarterly Journal of Mathematics, vol. 41 (1910), p. 170.

†I. Schur, Crelle's Journal, vol. 132 (1907), p. 113.

We proceed to determine some properties of representation groups of a non-cyclic abelian group. We shall especially study the representation group which is characterized by the fact that a set of independent generators of G' may be so chosen that to each one of these generators there corresponds at least one operator of the same order in G. Such a representation group will be called the *principal representation group* of the abelian group G'. It follows directly from the general theory of isomorphisms that there is only one such group.* For instance, the principal representation group of the four-group is the dihedral group of order 8, and the principal representation group of every other non-cyclic group of order p^2 is the non-abelian group of order p^3 which involves no operator of order p^2.

Every metabelian group whose group of inner isomorphisms is of odd order is conformal with an abelian group;† that is, a $(1, 1)$ correspondence can be established between the operators of such a metabelian group and those of an abelian group so that two corresponding operators have always the same order. In particular, every metabelian group of odd order is conformal with some abelian group. The fact that *every representation group of any abelian group of odd order is conformal with some abelian group* is evidently a special case of this theorem. The orders of the operators of such a representation group are therefore completely determined by the invariants of the conformal abelian group.

If G' is an abelian group of order p^m, p being any odd prime number, and if the invariants of G' are $p^{a_1}, p^{a_2}, \cdots, p^{a_\lambda}$, then the principal representation group of G' is conformal with the abelian group which has λ invariants each of which is equal to p^{a_1}, $\lambda - 1$ each equal to $p^{a_2}, \cdots$, one equal to p^{a_λ}. The total number of the invariants of this abelian group is therefore equal to $\frac{1}{2}\lambda(\lambda + 1)$. This is the largest number of invariants that an abelian group which is conformal with a representation group of G' can have. It is however possible that more than one of the representation groups are conformal with abelian groups having this number of invariants. The smallest number of invariants that an abelian group which is conformal with a representation group of G' can have is evidently $\frac{1}{2}\lambda(\lambda + 1) - \lambda + 1$. When G' is the non-cyclic group of order p^2, p being odd, this number is 2, as it evidently should be according to this formula.

When G' is the abelian group of order 2^m and of type $(1, 1, \cdots)$, the principal representation group of G' contains evidently exactly

$$-1 + (m + 1) \cdot 2^{\frac{1}{2}m(m-1)}$$

*Bulletin of the American Mathematical Society, vol. 3 (1897), p. 218.

†Bulletin of the American Mathematical Society, vol. 2 (1896), p. 140; vol. 6 (1900), p. 337.

operators of order 2. A necessary and sufficient condition that this representation group is conformal with an abelian group is that $m + 1$ is a power of 2. The representation groups of an abelian group of order 2^m are therefore sometimes, but not always, conformal with abelian groups. It may be observed that when a set of independent generators of the abelian group G' of order 2^m and of type $(1, 1, 1, \cdots)$ has been so chosen that each of these generators corresponds to an operator of order 2 in the principal representation group of G', then every other operator of order 2 in G' will correspond to operators of order 4 in this representation group. It is clearly always possible to find at least one representation group of this abelian group such that all of its operators, besides those in the commutator subgroup, are of order 4. The quaternion group is the simplest illustrative example of these representation groups.

The maximal order of an operator in the representation group of an abelian group G' is evidently equal to the product of the orders of the largest operator in M and the largest operator of G'. Moreover, a representation group of G' can always be constructed so that to each one of a set of independent generators of G' there corresponds an operator whose order is equal to the order of this generator multiplied by the order of an arbitrary operator in M. In fact, each of the operators which correspond to a set of independent generators of G' can be so selected as to generate an arbitrary operator of M. As a special case of this we have the theorem: *The principal representation group of G' contains no operator whose order exceeds that of the largest operator in G' when the order of G' is odd. When the order of G' is even, the order of the largest operator in its principal representation group cannot exceed twice that of the largest operator in G'.* This maximum order is attained only when the Sylow subgroup of order 2^m in G' has at least two equal largest invariants.

As a very special but interesting result from what precedes we may note the elementary fact, stated in the Introduction, that every possible non-cyclic abelian group has more than one representation group and that each of these representation groups is metabelian. In fact, such a non-cyclic group has a non-cyclic Sylow subgroup, and the principal representation group of this subgroup does not involve any operator whose order is equal to that of the largest operator which occurs in some one of the representation groups, unless the invariants of this Sylow subgroup are all equal to 2. In this special case we have observed that the number of distinct representation groups is at least equal to 2. A number of theorems relating to the possible number of different representation groups of a given abstract group were established by I. SCHUR in the memoirs to which we referred in the Introduction.

The representation groups of abelian groups furnish many simple illustrations of groups whose commutator subgroups involve operators which are not

commutators. W. B. FITE gave one example of such a group.* A very elementary infinite system of such groups may be established by observing that the representation groups of the abelian group of order p^m and of type $(1, 1, 1, \cdots)$ have a commutator subgroup of order $p^{\frac{1}{2}m(m-1)}$ while it has been observed that the number of their distinct commutators is less than p^{2m}. That is, *the commutator subgroup of a representation group of the abelian group of order p^m and of type $(1, 1, 1, \cdots)$ contains operators which are not commutators whenever $m > 4$.* As this theorem is true for every value of the prime number p and every value of $m > 4$, it establishes the existence of a doubly infinite system of such groups. It is also very easy to establish such systems in the case when the abelian group is not of type $(1, 1, 1, \cdots)$.

§ 4. *Representation groups of non-abelian groups of order p^m.*

If G' is a non-abelian group of order p^m, p being a prime number, it contains at least one series of invariant subgroups $H_1, H_2, \cdots, H_{m-1}$ of order $p, p^2, \cdots, p^{m-1}$ respectively, such that each of them is contained in all those which follow it. If one element of a commutator is in H_α, $\alpha = 1, 2, \cdots, m-1$, the commutator is found in $H_{\alpha-1}$, and if both of these elements are in H_α the commutator is in $H_{\alpha-2}$, and the same commutator results from an element in $H_{\alpha-1}$ and an element in H_α. In particular, all the possible commutators of G' must be contained in H_{m-2}, as is well known.

Suppose that G' is the central quotient group of a group K. The commutators of K which have one element in the subgroup corresponding to H_1 must have orders which divide p, according to § 2. All the operators of order p in the subgroup corresponding to the identity of K clearly generate a subgroup which involves no operator of order p^2. Similarly, the commutators of K which have one element in the subgroup corresponding to H_2 must have orders which divide p^2, and all the operators whose orders divide p^2, in the subgroup of K which corresponds to H_1, generate a subgroup which involves no operator of order p^3.

We proceed to prove, by complete induction, that all the commutators of K, which have one element in the subgroup of K which corresponds to H_α, have orders which divide p^α, and all the operators of K whose orders divide p^α and which correspond to $H_{\alpha-1}$ generate a group involving no operator of order $p^{\alpha+1}$. In fact, if s_α is any operator of K contained in the subgroup corresponding to H_α and if t is any other operator of K, we have the equation:

$$(t^{-1} s_\alpha t)^p = t^{-1} s_\alpha^p t = (C_\alpha s_\alpha)^p = C_\alpha s_\alpha C_\alpha s_\alpha^{-1} \cdot s_\alpha^2 C_\alpha s_\alpha^{-2} \cdot \cdots \cdot s_\alpha^{p-1} C_\alpha s_\alpha^{1-p} \cdot s_\alpha^p.$$

* These Transactions, vol. 3 (1902), p. 340.

Hence

$$C_{a-1}\, s_a^p = C_a\, C_{a-1}^{(1)}\, C_a C_{a-1}^{(2)}\, C_a \cdot \cdots\, C_{a-1}^{(p-1)}\, C_a \cdot s_a^p$$

$$C_{a-1} = C_a C_{a-1}^{(1)} C_a^{-1} \cdot C_a^2\, C_{a-1}^{(2)}\, C_a^{-2} \cdot \cdots\, C_a^{p-1}\, C_{a-1}^{(p-1)}\, C_a^{1-p} \cdot C_a^p .$$

All the operators between periods, except possibly the last, in this final equation are commutators which correspond to H_{a-2}. Hence we may assume that they together with C_{a-1} cannot generate any operator whose order exceeds p^{a-1}. This proves that the order of C_a cannot exceed p^a.

If t_1, t_2 are two operators of K which correspond to H_{a-1} and whose orders divide p^a, the order of their product must also divide p^a as may be seen from the following equations

$$\begin{aligned}(t_1 t_2)^p &= t_1 t_2 t_1 t_2^{-1} t_2^2 t_1 t_2^{-2} \cdots t_2^{p-1} t_1 t_2^{1-p} \cdot t_2^p \\ &= t_1\, C_{a-2}^{(1)}\, t_1\, C_{a-2}^{(2)}\, t_1 \cdots C_{a-2}^{(p-1)}\, t_1 \cdot t_2^p \\ &= C_{a-2}\, t_1^p\, t_2^p .\end{aligned}$$

As each of these three operators has an order which divides p^{a-1}, and as it corresponds to H_{a-2}, it has been proved, by complate induction, that *all the commutators of K involving one element which corresponds to H_a must have orders which divide p^a*. As a special case we have the theorem: *The order of every representation group of a group of order p^m, p being any prime number, is of the form p^m, and the orders of all its commutators divide p^{m-1}.*

The preceding theorem is useful to find an upper limit for the order of a representation group of a group of order p^m. In particular, this order is always finite since both the orders of the commutators and the number of the distinct commutators are finite whenever m is finite. For the sake of illustration we shall consider the representation groups of the non-abelian group of order p^3. Not more than two independent generators of M can arise from commutators involving elements which correspond to the invariant operators of this abelian group, and not more than one additional independent generator of M can arise from the commutators whose elements correspond to the non-invariant operators of this non-abelian group of order p^3. As the orders of all of these independent generators divide p, it results that the order of a representation group of a non-abelian group of order p^3 cannot exceed p^6. It is evident that the non-abelian group of order p^3 which is conformal with the abelian group of type $(1, 1, 1)$ has a representation group of order p^6, for all odd values of p; while the non-abelian group which is conformal with the abelian group of type $(2, 1)$ is its own representation group for all odd values of p.*

In general, it is evident from the theorem stated above that the subgroup

* These Transactions, vol. 10 (1909), p. 395.

M, which is called the *multiplier* of G', can not have more than $m-1$ independent generators arising from commutators which have an element corresponding to H_1 in G', and that the orders of all these independent generators divide p. Similarly, there cannot be more than $m-2$ additional generators in m, arising from commutators involving elements which correspond to H_2, and the orders of each of these divides p^2, etc. As the orders of these independent generators cannot always have this upper limit, since G' is non-abelian, it results directly that the order of M is always less than p^k, where

$$k = m-1+2(m-2)+3(m-3)\cdots+m-1.$$

In the special case considered in the preceding paragraph the order of M has its maximal value according to this formula, but it is evident that M will generally have a much lower order.

In conclusion it seems desirable to direct attention to a very elementary general theorem which is illustrated by the representation groups of the special categories of groups considered above; viz., the theorem that *if a group G' has more than one representation group all of these representation groups have the same commutator subgroup.* In fact, if they did not have the same commutator subgroup we could establish an isomorphism between them with respect to their common quotient group G' and thus obtain a larger representation group of G'. As special cases of this theorem we have the results that in all the representation groups of G' the multipliers are simply isomorphic, and if a group is generated by its commutators it has only one representation group. Both of these theorems were explicitly stated by I. Schur in the memoirs to which we referred in the Introduction.

ON THE ACCURACY OF TRIGONOMETRIC INTERPOLATION*

BY

DUNHAM JACKSON

Let $f(x)$ be a function of x, with the period 2π, which takes on known values, determined by physical measurement or otherwise, at $2n+1$ points equally spaced throughout a period. It is well known that a finite trigonometric sum, of the nth order at most, which coincides with $f(x)$ at the points in question, is given by formulæ analogous to those which characterize the Fourier's development of $f(x)$, the integrals of the latter development being replaced by finite sums involving the known quantities. It has been found by FABER that in the case of continuous functions the ordinary sufficient conditions, in the way of restrictions on $f(x)$, for the convergence of the Fourier's series are sufficient also to insure the convergence of the interpolating function to the value $f(x)$ at all points, as the number n is indefinitely increased; although in cases not covered by these explicit conditions one development may converge while the other does not.† In the present note it is shown that a method by which the author has studied the rapidity of convergence of Fourier's series is adapted to the treatment of the corresponding problem in interpolation,‡ and yields similar results. In fact, these results are obtained by a simple combination of materials already at hand. It is further pointed out how a finite number of observed values may be used to define a formula of approximation which converges more rapidly in certain cases than the ordinary interpolation formula.

I. ORDINARY INTERPOLATION.

Let $f(x)$ be a function of x having the period 2π. Let $x_1, x_2, \cdots, x_{2n+1}$ be a set of values of x such that

$$x_{i+1} - x_i = \frac{2\pi}{2n+1} \qquad (i = 1, 2, \cdots, 2n).$$

* Presented to the Society, February 22, 1913.

† FABER, Mathematische Annalen, vol. 69 (1910), pp. 372–443; see pp. 417–443. At the close of the article, FABER considers also certain classes of discontinuous functions. The sufficiency of the condition of limited variation had previously been deduced by DE LA VALLÉE POUSSIN, Bulletins de l'Académie royale de Belgique, Classe des Sciences, 1908, pp. 319–403.

‡ DE LA VALLÉE POUSSIN (loc. cit., p. 389) obtains the result that if $f(x)$ possesses a derivative having limited variation, the error of the interpolation formula nowhere exceeds a constant multiple of $1/n$. This theorem is not among those obtained in the present paper.

If the coefficients in the trigonometric sum

$$S_n(x) = \frac{a_0}{2} + a_1 \cos x + \cdots + a_n \cos nx + b_1 \sin x + \cdots + b_n \sin nx$$

are determined by the formulæ

$$a_h = \frac{2}{2n+1} \sum_{i=1}^{2n+1} f(x_i) \cos hx_i, \qquad b_h = \frac{2}{2n+1} \sum_{i=1}^{2n+1} f(x_i) \sin hx_i \qquad (h = 0, 1, \cdots, n),$$

the sum can be expressed in the form*

$$S_n(x) = \frac{1}{2n+1} \sum_{i=1}^{2n+1} f(x_i) \frac{\sin[(n+\frac{1}{2})(x_i - x)]}{\sin \frac{1}{2}(x_i - x)},$$

by means of a familiar identity, and from this expression it appears that $S_n(x_j) = f(x_j)$ at each of the $2n+1$ points x_j, all the terms but one under the sign of summation being zero. It is the only trigonometric sum of order n or lower which has this property, since two such sums which are equal for more than $2n$ distinct values (mod 2π) of the argument must be identical.

It was proved by FABER† that if $f(x)$ never exceeds 1 in absolute value, then the absolute value of $S_n(x)$ can not exceed a certain constant multiple of $\log n$. The proof, which may be given here for the sake of completeness, is as follows. With the hypothesis made concerning $f(x)$,

$$|S_n(x)| \leqq \frac{1}{2n+1} \sum_{i=1}^{2n+1} \left| \frac{\sin[(n+\frac{1}{2})(x_i - x)]}{\sin \frac{1}{2}(x_i - x)} \right|.$$

Set $\frac{1}{2}(x_i - x) = u_i$; then

$$|S_n(x)| \leqq \frac{1}{2n+1} \sum_{i=1}^{2n+1} \left| \frac{\sin(2n+1)u_i}{\sin u_i} \right|.$$

In view of the periodicity of the quotient, it may be assumed without loss of generality that all the variables u_i lie in the interval from $-\frac{1}{2}\pi$ to $\frac{1}{2}\pi$. Let v_0 be the smallest of the numbers $|u_i|$ (or one of the two smallest, if two are equal), v_1 the second, and so on. Then we have, on rearranging the sum,

$$|S_n(x)| \leqq \sum_{i=0}^{2n} \left| \frac{\sin(2n+1)v_i}{(2n+1)\sin v_i} \right|,$$

where, as is readily seen,

$$\frac{i\pi}{2(2n+1)} \leqq v_i \leqq \frac{(i+1)\pi}{2(2n+1)}.$$

Now

* Throughout this paper, an expression which, regarded as a quotient, becomes meaningless for certain values of the variable by taking on the form $0/0$, will be understood to be defined for those values so as to be a continuous function.

† Loc. cit., pp. 419–420.

$$(1) \qquad \left|\frac{\sin(2n+1)v_0}{(2n+1)\sin v_0}\right| = \left|\frac{\sin(2n+1)v_0}{(2n+1)v_0}\cdot\frac{v_0}{\sin v_0}\right| \leqq 1,$$

since $(\sin v)/v$ is a monotonically decreasing function in the interval $(0, \pi)$ which includes both the points v_0 and $(2n+1)v_0$. Similarly, the absolute value of the term in v_1 does not exceed 1. For $i \geqq 2$,

$$\left|\frac{\sin(2n+1)v_i}{(2n+1)\sin v_i}\right| \leqq \frac{1}{(2n+1)\sin v_i};$$

since $\sin v \geqq 2v/\pi$ for $0 \leqq v \leqq \frac{1}{2}\pi$, we may write

$$(2) \qquad (2n+1)\sin v_i \geqq (2n+1)\sin\frac{i\pi}{2(2n+1)} \geqq i,$$

and hence

$$\sum_{i=2}^{2n}\left|\frac{\sin(2n+1)v_i}{(2n+1)\sin v_i}\right| \leqq \sum_{i=2}^{2n}\frac{1}{i} < \log 2n.$$

Consequently

$$|S_n(x)| < 2 + \log 2 + \log n < 5\log n \qquad (n \geqq 2).$$

It would be easy to replace the constant 5 by a smaller value, but we shall not make any effort to do so here.

From the inequality just obtained, it follows immediately that if $|f(x)| \leqq \epsilon$, then $|S_n(x)| \leqq 5\epsilon\log n$, and $|f(x) - S_n(x)| \leqq \epsilon + 5\epsilon\log n \leqq 7\epsilon\log n$, $n \geqq 2$.

Now let the restriction on $f(x)$ be replaced by the broader hypothesis that there exists a trigonometric sum $T_n(x)$, of the nth order at most, such that for all values of x we have the inequality

$$|f(x) - T_n(x)| \leqq \epsilon.$$

The interpolating function $S_n(x)$ that corresponds to $f(x)$ may be obtained by adding those for $T_n(x)$ and for $f(x) - T_n(x)$. The former is identical with $T_n(x)$, being equal to it at $2n+1$ points, and so an error is committed only in approximating to the difference $f(x) - T_n(x)$, which does not exceed ϵ. It follows that the error can never be greater than $7\epsilon\log n$ if $n \geqq 2$, and we have, to sum up,*

THEOREM I. *If $f(x)$ can be approximately represented by a finite trigonometric sum of the nth order at most, $n \geqq 2$, with an error never exceeding ϵ, then $f(x)$ is represented by its interpolating function $S_n(x)$ with an error not exceeding $7\epsilon\log n$.*

This can be immediately combined with theorems of the author concerning the accuracy of approximation by trigonometric sums† to give a variety of more explicit results. To mention only two:

* Cf. LEBESGUE, *Sur les intégrales singulières*, Annales de la Faculté de Toulouse, ser. 3, vol. 1 (1910), pp. 25–117, for the corresponding result in the case of Fourier's series (see pp. 116–117).

† These Transactions, vol. 13 (1912), pp. 491–515. This article will be referred to hereafter as A. Also, these Transactions, pp. 343–364 of the present volume.

If $\lim_{\delta=0}[f(x+\delta)-f(x)]\log|\delta| = 0$, uniformly for all values of x, then $S_n(x)$ uniformly approaches $f(x)$ as its limit when n is indefinitely increased.*

If $f(x)$ everywhere satisfies the Lipschitz condition†

$$|f(x'')-f(x')| \leqq \lambda |x''-x'|,$$

then, for all values of x, and for $n \geqq 2$,

$$|f(x)-S_n(x)| \leqq \frac{21\lambda \log n}{n}.$$

The preceding work also makes it easy to estimate the possible effect of errors‡ in the determination of the quantities $f(x_i)$. Let $S_n(x)$ be the interpolating function defined by the true values $f(x_i)$, and $\bar{S}_n(x)$ the function defined by a set of observed values $\bar{f}(x_i)$. The difference between these two functions is the corresponding function defined by the differences $f(x_i)-\bar{f}(x_i)$, and, if the greatest of the latter does not exceed ϵ, the maximum of $|S_n(x)-\bar{S}_n(x)|$ can not exceed $5\epsilon \log n$. It would be possible to replace the coefficient 5 by a smaller one, as has been said before, but it is interesting to remark that it would not be possible to dispense with the factor $\log n$, or even to replace it by one which becomes infinite less rapidly with n, since, as FABER shows, the maximum of $|S_n(x)|$ for functions $f(x)$ not exceeding 1 in absolute value actually is of the order§ of $\log n$.

II. A MODIFIED FORMULA.

For comparison with $S_n(x)$, let us define a function $\Sigma_m(x)$ as follows:

$$\Sigma_m(x) = H_m \sum_{i=1}^{2m} f(x_i)\left[\frac{\sin \tfrac{1}{2}m(x_i-x)}{m \sin \tfrac{1}{2}(x_i-x)}\right]^4, \tag{3}$$

where

$$\frac{1}{H_m} = \sum_{i=1}^{2m}\left[\frac{\sin \tfrac{1}{2}m(x_i-x)}{m \sin \tfrac{1}{2}(x_i-x)}\right]^4.$$

The points x_i are assumed to be such that‖

* This was deduced by FABER (loc. cit., p. 422); the particular theorem used here concerning approximation by trigonometric sums was proved by LEBESGUE, loc. cit., pp. 115–116.

† See A, Theorems I, VI.

‡ The fact that, in the case of polynomial interpolation with equidistant ordinates, errors of observation may produce a disproportionately large effect, was pointed out by DE LA VALLÉE POUSSIN, loc. cit., pp. 321–322; see also p. 346 of the same article.

§ FABER, loc. cit., p. 424. If $f(x_i) = (-1)^{i-1}$, and $x_1 - x = \pi/(2n+1)$, then

$$f(x_i)\sin[(n+\tfrac{1}{2})(x_i-x)] = +1$$

for all values of i, and it is readily seen that $S_n(x)$ is of the order of $\log n$.

‖ It would be possible, without essential change, to use instead an odd number, $2m-1$, of equally spaced points.

$$x_{i+1} - x_i = \frac{\pi}{m} \qquad (i = 1, 2, \cdots, 2m-1).$$

Let us first notice that the apparent dependence of H_m on the variable x is illusory; H_m really depends only on m. This may be seen as follows. The fourth power under the sign of summation can be expressed in the form*

$$A_{m0} + A_{m1} \cos(x_i - x) + A_{m2} \cos 2(x_i - x) + \cdots + A_{m,\,2m-2} \cos[(2m-2)(x_i - x)] = \sum_{h=0}^{2m-2} A_{mh} (\cos hx_i \cos hx + \sin hx_i \sin hx),$$

the coefficients A_{mh} being constants. It is a well-known property of the trigonometric functions that if $y_1, \cdots, y_q$ are q points disposed at successive intervals of $2\pi/q$, and p is not divisible by q, then†

$$\sum_{i=1}^{q} \cos py_i = \sum_{i=1}^{q} \sin py_i = 0.$$

For the application that we wish to make, q is $2m$, and p is one of the numbers $0, 1, \cdots, 2m-2$, and so the sum defining $1/H_m$ reduces to $2mA_{m0}$.

From this fact and the expression used above for the fourth power involved, it appears that $\Sigma_m(x)$ is a trigonometric sum in x, of order not higher than $2(m-1)$. It is a formula of interpolation, in the sense that it is completely determined by a finite number of values of the function represented, though it does not, like $S_n(x)$, become equal to $f(x)$ at the points x_i, in spite of the fact that it is of higher order than $S_n(x)$ in proportion to the number of points used. We shall proceed to establish the property which constitutes its chief claim to attention: If $f(x)$ satisfies a Lipschitz condition, it is represented by $\Sigma_m(x)$ with an error which approaches zero at least as fast as $1/n$ in order of magnitude, if we denote by n the maximum order $2m-2$ of the trigonometric sum.

Suppose that $f(x)$ satisfies everywhere the Lipschitz condition

$$|f(x'') - f(x')| \leqq \lambda |x'' - x'|.$$

* See A, p. 493. The absence of sine-terms follows from the fact that the expression is an even function of $x_i - x$.

† Cf. e. g. BÔCHER, *Introduction to the theory of Fourier's series*, Annals of Mathematics, ser. 2, vol. 7 (1906), pp. 81–152; p. 135. For an analytic proof, a starting point is the fact that when $p = 1$ and one of the y_i's is zero, $\Sigma \cos y_i$ and $\Sigma \sin y_i$ are the real and imaginary parts respectively of the sum of the qth roots of unity, which is zero if $q > 1$. Then the cases may be considered successively that p is any number relatively prime to q, and that p and q have a greatest common divisor greater than 1 but less than q, still with the assumption that one of the y_i's is zero. The last generalization may be effected by letting $y_i = y_1 + y_i'$, where y_1 is arbitrary.

We may write

$$\Sigma_m(x) - f(x) = H_m \sum_{i=1}^{2m} [f(x_i) - f(x)] \left[\frac{\sin \frac{1}{2}m(x_i - x)}{m \sin \frac{1}{2}(x_i - x)}\right]^4, \tag{4}$$

or, setting $\frac{1}{2}(x_i - x) = u_i$,

$$|\Sigma_m(x) - f(x)| = \left| H_m \sum_{i=1}^{2m} [f(x + 2u_i) - f(x)] \left[\frac{\sin mu_i}{m \sin u_i}\right]^4 \right| \tag{5}$$

$$\leqq 2\lambda H_m \sum_{i=1}^{2m} |u_i| \left[\frac{\sin mu_i}{m \sin u_i}\right]^4.$$

Because of the periodicity of the functions involved in the first line of this last relation, we may assume without loss of generality that all the variables u_i there, and so also in the following line, are in the interval from $-\frac{1}{2}\pi$ to $+\frac{1}{2}\pi$.

The work of determining approximately the magnitude of H_m has essentially been performed in the article A. For, if we recall the notation J_m of that article,* we have

$$J_m = m \int_0^{\pi/2} \left[\frac{\sin mu}{m \sin u}\right]^4 du = m \int_0^{\pi/2} A_{m0}\, du = \frac{m\pi}{2} A_{m0} = \frac{\pi}{4} \cdot \frac{1}{H_m}.$$

As it was shown that $\frac{1}{2}\pi \geqq J_m > \frac{1}{3}\pi$, for all values of m, it follows that $\frac{1}{2} \leqq H_m < \frac{3}{4}$.

In the sum which still remains to be evaluated, let the smallest of the quantities $|u_i|$ (or, if this is indeterminate, one of the two smallest) be set equal to v_0, the next to v_1, and so on. Then

$$\sum_{i=1}^{2m} |u_i| \left[\frac{\sin mu_i}{m \sin u_i}\right]^4 = \sum_{i=0}^{2m-1} v_i \left[\frac{\sin mv_i}{m \sin v_i}\right]^4,$$

where

$$\frac{i\pi}{4m} \leqq v_i \leqq \frac{(i+1)\pi}{4m}.$$

Adapting the relation (1) to the present case, we see that

$$\left[\frac{\sin mv_0}{m \sin v_0}\right]^4 \leqq 1, \qquad \left[\frac{\sin mv_1}{m \sin v_1}\right]^4 < 1.$$

For $i \geqq 2$, by reasoning similar to that employed in connection with (2),

$$m \sin v_i \geqq m \sin \frac{i\pi}{4m} \geqq \frac{i}{2},$$

and

$$\left[\frac{\sin mv_i}{m \sin v_i}\right]^4 \leqq \left(\frac{2}{i}\right)^4.$$

Hence

* A, pp. 503–507.

$$\sum_{i=0}^{2m-1} v_i \left[\frac{\sin mv_i}{m \sin v_i}\right]^4 \leqq v_0 + v_1 + \sum_{i=2}^{2m-1} v_i \left(\frac{2}{i}\right)^4.$$

Now $v_0 + v_1 = \pi / (2m)$, as is seen by reference to the definition of the symbols v_i; and

$$\sum_{i=2}^{2m-1} v_i \left(\frac{2}{i}\right)^4 \leqq \sum_{i=2}^{2m-1} \frac{(i+1)\pi}{4m} \left(\frac{2}{i}\right)^4 = \frac{4\pi}{m} \sum_{i=2}^{2m-1} \frac{i+1}{i^4} < \frac{4\pi}{m}\left(\int_1^\infty \frac{du}{u^3} + \int_1^\infty \frac{du}{u^4}\right)$$

$$= \frac{4\pi}{m}\left(\frac{1}{2} + \frac{1}{3}\right) = \frac{10\pi}{3m}.$$

Consequently

$$\sum_{i=0}^{2m-1} v_i \left[\frac{\sin mv_i}{m \sin v_i}\right]^4 < \frac{\pi}{m}\left(\frac{1}{2} + \frac{10}{3}\right) = \frac{23\pi}{6m},$$

and so, by (5) and the inequality for H_m,

$$| \Sigma_m(x) - f(x) | \leqq \frac{23\pi}{4} \cdot \frac{\lambda}{m}.$$

If we denote by n, as was suggested before, the maximum order of $\Sigma_m(x)$, which is less than $2m$, we have finally:

THEOREM II. *If the function $f(x)$, of period 2π, everywhere satisfies the Lipschitz condition $|f(x'') - f(x')| \leqq \lambda |x'' - x'|$, it is approximately represented by the finite trigonometric sum $\Sigma_m(x)$, of order not higher than n, with an error not exceeding $37\lambda / n$.*

The coefficient 37 is not particularly significant; no attempt has been made to reduce it to as small a value as possible. The main point is that we have found here an upper limit for the error which is of higher order in its approach to zero than that which we had for the ordinary interpolation formula. We have not proved, to be sure, that the order of the limit of error which we found in the earlier case is the best that can be obtained, but it is not difficult to show, by the method applied by LEBESGUE* to the corresponding problem in the case of Fourier's series, that the error of the approximating function $S_n(x)$ for a function $f(x)$ satisfying a Lipschitz condition can actually exceed a constant multiple of $(\log n) / n$ for infinitely many values of n. The expression $\Sigma_m(x)$ accordingly is really superior to the other in the order of its approximation to certain functions of the class considered.

It may be observed further that errors in the quantities $f(x_i)$, none of which exceeds ϵ in absolute value, can not affect $\Sigma_m(x)$ at any point by more than ϵ, as compared with a quantity of the order of $\epsilon \log n$ in the case of $S_n(x)$, and that $\Sigma_m(x)$ converges uniformly to the value $f(x)$, as m is indefinitely

* LEBESGUE, Bulletin de la Soc. Math. de France, vol. 38 (1910), pp. 184–210; pp. 203–206.

increased, for every continuous function $f(x)$. The former assertion is coextensive with the statement that $|\Sigma_m(x)|$ is never greater than 1 if $|f(x)|$ never exceeds 1, a fact which is obvious from the definition of $\Sigma_m(x)$.

To prove the convergence property suppose that ϵ is an arbitrary positive quantity. Let δ be a positive number such that $|f(x'') - f(x')| \leqq \frac{1}{2}\epsilon$ whenever $|x'' - x'| < \delta$; the existence of δ is a consequence of the hypothesis that $f(x)$ is everywhere continuous. Let x have any particular value, and divide the sum on the right hand side of (4) into two parts, one containing the terms* for which $|x_i - x| \leqq \delta$, and the other, the remaining terms. The first part does not exceed $\frac{1}{2}\epsilon$ in absolute value, for any value of m, in consequence of the definition of δ. In the second part of the sum, $|f(x_i) - f(x)|$ does not exceed twice the maximum of $|f(x)|$, and $\sin \frac{1}{2}(x_i - x) \geqq \sin \frac{1}{2}\delta$; consequently each term of the second part is inferior in absolute value to $1/m^4$ multiplied by a quantity independent of m and x, and the whole second part does not exceed a constant multiple of $1/m^3$. As soon as m is greater than a suitable quantity, independent of x, the second part too becomes smaller than $\frac{1}{2}\epsilon$, and the uniform convergence is established.

The properties of $\Sigma_m(x)$ last developed, while interesting in themselves, are less characteristic than the one on which the principal stress has been laid; they are shared, for example, by the simpler expression which is obtained if the fourth power in the definition of $\Sigma_m(x)$ is replaced by the corresponding square.

To apply the formula $\Sigma_m(x)$ to the problem of polynomial interpolation, the given ordinates, to be sure, being unequally spaced, suppose that $g(y)$ is a function of y defined in the interval $-1 \leqq y \leqq 1$. Suppose that the values of $g(y)$ are known, either at the m points

$$y_i = \cos\frac{(2i-1)\pi}{2m} \qquad (i = 1, 2, \cdots, m),$$

or at the $m+1$ points

$$y_i = \cos\frac{i\pi}{m} \qquad (i = 0, 1, 2, \cdots, m).$$

If we set $y = \cos x$, $g(y) = g(\cos x) = f(x)$, we have the values of $f(x)$ given for a set of values of x symmetrically situated with respect to the origin, and equally spaced at intervals of π/m. Let those which lie in the interval $-\pi < x \leqq \pi$ be used as the points x_i for defining a function $\Sigma_m(x)$ corresponding to our present function $f(x)$. If we replace x by $-x$ in this expression, and at the same time replace x_i by $-x_i$, which amounts only to

* We suppose the numbers x_i replaced, if necessary, by values congruent to them (mod 2π), so that none of the differences $|x_i - x|$ is greater than π.

a rearrangement of the terms in the sum (3), because of the symmetrical arrangement of the points x_i, we see that $\Sigma_m(x)$ remains unchanged. Being an even function, it is a polynomial in $\cos x$, of degree $2(m-1)$ at most. We have thus a means of approximating to $g(y)$ by a polynomial in y. The properties of the trigonometric interpolation formula developed above can be immediately interpreted with reference to the present case.*

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

* Cf. A, p. 495.

ON A SIMPLE TYPE OF IRREGULAR SINGULAR POINT*

BY

GEORGE D. BIRKHOFF

Introduction.

A simple type of irregular singular point is furnished by the linear differential equation of the second order

$$\frac{d^2y}{dx^2} + p(x)\frac{dy}{dx} + q(x)\,y = 0 \tag{1}$$

in which the coefficients $p(x)$ and $q(x)$ are analytic at $x = \infty$. The point $x = \infty$ is then an irregular singular point of rank 1 unless $xp(x)$ and $x^2 q(x)$ remain finite at infinity. If we write $\lim_{x=\infty} p(x) = p_0$, $\lim_{x=\infty} q(x) = q_0$, and if the roots b_1, b_2 of the quadratic equation

$$b^2 + p_0 b + q_0 = 0$$

are distinct, the transformation of variables from x, y to $\bar{x}$, $\bar{y}$,

$$y = e^{b_1 x}\, x^{c_1}\, \bar{y}, \qquad x = \bar{x}\,(b_2 - b_1),$$

will take (1) into another equation of the same form, in which however

$$p(x) = -1 + \frac{p_1}{x} + \cdots, \qquad q(x) = \frac{q_2}{x^2} + \frac{q_3}{x^3} + \cdots, \tag{2}$$

provided that a proper determination of the constant c_1 be made.

In the present paper I wish to consider the solutions of (1) in the vicinity of $x = \infty$, but under the restriction that $p(x)$ and $q(x)$ have the form (2).† The method of attack is essentially the same as that which I have employed earlier in the consideration of singular points.‡ In this special case the results may be given a more striking form. The last two theorems have no analoguesin my earlier paper and are susceptible of wide generalization.

---

* Presented to the Society, September 12, 1911.

† This type of equation has been considered in a paper by HORN, Journal für Mathematik, vol. 133 (1908), pp. 19–67.

‡ These Transactions, vol. 10 (1909), pp. 436–470.

The treatment will be based on the following auxiliary theorem:* *Let* $l_{11}(x)$, $l_{12}(x)$, $l_{21}(x)$, $l_{22}(x)$ *be functions single-valued and analytic for* $|x| \geqq R$ *(but not necessarily at* $x = \infty$*) and such that* $l_{11}(x)\, l_{22}(x) - l_{12}(x)\, l_{21}(x)$ *does not vanish for* $|x| \geqq R$. *Then there exists a set of functions* $a_{11}(x)$, $a_{12}(x)$, $a_{21}(x)$, $a_{22}(x)$ *analytic at* $x = \infty$ *and reducing to* 1, 0, 0, 1 *respectively at* $x = \infty$, *and a set of entire functions* $e_{11}(x)$, $e_{12}(x)$, $e_{21}(x)$, $e_{22}(x)$ *for which* $e_{11}(x)\, e_{22}(x) - e_{12}(x)\, e_{21}(x)$ *is nowhere zero in the finite plane, such that*

$$
\begin{aligned}
l_{11}(x) &= [\, a_{11}(x)\, e_{11}(x) + a_{12}(x)\, e_{21}(x)\,]\, x^{k_1}, \\
l_{12}(x) &= [\, a_{11}(x)\, e_{12}(x) + a_{12}(x)\, e_{22}(x)\,]\, x^{k_2}, \\
l_{21}(x) &= [\, a_{21}(x)\, e_{11}(x) + a_{22}(x)\, e_{21}(x)\,]\, x^{k_1}, \\
l_{22}(x) &= [\, a_{21}(x)\, e_{12}(x) + a_{22}(x)\, e_{22}(x)\,]\, x^{k_2},
\end{aligned} \tag{3}
$$

where k_1 *and* k_2 *are integers.*

§ 1. *The Canonical Form of Equation.*

If we set $x\,(dy/dx) = z$, we may replace the single equation (1) by the linear differential system

$$
\frac{dy}{dx} = \frac{1}{x} z, \qquad \frac{dz}{dx} = -\, xq(x)\, y + \left(-\, p(x) + \frac{1}{x} \right) z. \tag{4}
$$

Let us now make a linear transformation of the dependent variables

$$
y = a_{11}(x)\, Y + a_{12}(x)\, Z, \qquad z = a_{21}(x)\, Y + a_{22}(x)\, Z, \tag{5}
$$

where the functions $a_{11}(x)$, $a_{12}(x)$, $a_{21}(x)$, $a_{22}(x)$ are as yet unspecified functions for which $\Delta(x) = a_{11}(x)\, a_{22}(x) - a_{12}(x)\, a_{21}(x)$ is not identically zero. The new linear differential system in Y, Z will be

$$
\frac{dY}{dx} = P_{11}(x)\, Y + P_{12}(x)\, Z, \qquad \frac{dZ}{dx} = P_{21}(x)\, Y + P_{22}(x)\, Z, \tag{6}
$$

where the coefficients of Y and Z have the values

* This theorem is the special case $n = 2$ of a theorem which I have proved, Mathematische Annalen, vol. 74 (1913), pp. 122–133. This proof, based on the theory of linear integral equations, was first presented by me in 1911 (see Bulletin of the American Mathematical Society, vol. 18, p. 64) although I was in possession of it a year or two earlier; a proof by an entirely different method was given by me in 1908 (see these Transactions, vol. 10 (1909), p. 438). I have recently noticed that the theorem is a special case of a very important theorem due to Hilbert, Göttinger Nachrichten (1905), pp. 307–330, and Plemelj, Monatshefte für Mathematik und Physik, vol. 19 (1908), pp. 211–222. This relationship becomes obvious if one identifies the functions $l_{ij}(x)$ with the functions $\alpha_{ij}(x)$ in Plemelj's paper, and solves the equations like (3) for $a_{ij}(x)$; for the boundary curve (*Randkurve*) may be taken the curve $|x| = R$. Furthermore the method used in my second proof is essentially the same as that used by them.

$$
\begin{aligned}
P_{11} &= \frac{1}{\Delta}\left[a_{22}\left(\frac{1}{x}a_{21} - a'_{11}\right) - a_{12}\left(-xqa_{11} + \left(-p + \frac{1}{x}\right)a_{21} - a'_{21}\right)\right],\\
P_{12} &= \frac{1}{\Delta}\left[a_{22}\left(\frac{1}{x}a_{22} - a'_{12}\right) - a_{12}\left(-xqa_{12} + \left(-p + \frac{1}{x}\right)a_{22} - a'_{22}\right)\right],\\
P_{21} &= \frac{1}{\Delta}\left[-a_{21}\left(\frac{1}{x}a_{21} - a'_{11}\right) + a_{11}\left(-xqa_{11} + \left(-p + \frac{1}{x}\right)a_{21} - a'_{21}\right)\right],\\
P_{22} &= \frac{1}{\Delta}\left[-a_{21}\left(\frac{1}{x}a_{22} - a'_{12}\right) + a_{11}\left(-xqa_{12} + \left(-p + \frac{1}{x}\right)a_{22} - a'_{22}\right)\right].
\end{aligned}
\tag{7}
$$

The accent here denotes differentiation, and for brevity the letter x is suppressed in the functional notation. If the functions $a_{11}(x)$, $a_{12}(x)$, $a_{21}(x)$, $a_{22}(x)$ are analytic at $x = \infty$ and reduce to $1, 0, 0, 1$ respectively at $x = \infty$, the function $\Delta(x)$ reduces to 1 at $x = \infty$, and a direct computation by means of (7) gives

$$
\begin{aligned}
P_{11}(x) &= \frac{p_{11}^{(2)}}{x^2} + \frac{p_{11}^{(3)}}{x^3} + \cdots,\\
P_{12}(x) &= \frac{r}{x} + \frac{p_{12}^{(2)}}{x^2} + \cdots,\\
P_{21}(x) &= \frac{s}{x} + \frac{p_{21}^{(2)}}{x^2} + \cdots,\\
P_{22}(x) &= 1 + \frac{1 - p_1}{x} + \frac{p_{22}^{(2)}}{x^2} + \cdots.
\end{aligned}
\tag{8}
$$

Later we shall make use of these expansions in series.

By the aid of the auxiliary theorem above stated we shall choose the transformation (5) so that the equations (6) reduce to a very simple form.

Consider the solutions of (1) in the vicinity of $x = \infty$. According to the fundamental existence theorem for ordinary linear differential equations, every solution of (1) will be analytic for $|x| \geqq R$, if R is chosen so large that $p(x)$ and $q(x)$ are analytic for $|x| \geqq R$; and the general solution will be $c_1 y_1 + c_2 y_2$, where c_1, c_2 are arbitrary constants and y_1, y_2 are any pair of linearly independent particular solutions. But y_1 and y_2 are not in general single-valued for $|x| \geqq R$; in fact, when x makes a positive circuit of $x = \infty$, y_1 and y_2 are only known to be replaced by some pair $\bar{y}_1$, $\bar{y}_2$ of linearly independent solutions so that

$$
\bar{y}_1 = c_1 y_1 + c_2 y_2, \qquad \bar{y}_2 = d_1 y_1 + d_2 y_2, \qquad c_1 d_2 - c_2 d_1 \neq 0.
$$

Thus it is evident that y_1, y_2 undergo a linear transformation when x makes a positive circuit of $x = \infty$. According to the well-known facts concerning

such transformations, it will then be possible so to modify the choice of y_1, y_2 that we have one of the two cases:

$$(9) \qquad \bar{y}_1 = \rho_1 y_1, \qquad \bar{y}_2 = \rho_2 y_2; \quad \text{or} \quad \bar{y}_1 = \rho_1 y_1, \qquad \bar{y}_2 = \rho_1 y_1 + y_2.$$

The second of these alternatives is to be regarded as exceptional ($\rho_1 = \rho_2$), and we suppose at present that the first alternative is before us.

If we define λ_1 and λ_2 by means of the equations

$$\lambda_1 = \frac{-1}{2\pi\sqrt{-1}} \log \rho_1, \qquad \lambda_2 = \frac{-1}{2\pi\sqrt{-1}} \log \rho_2,$$

thus leaving λ_1 and λ_2 arbitrary up to additive integers, and then define $l_{11}(x), l_{12}(x), l_{21}(x), l_{22}(x)$ by means of the equations

$$(10) \qquad \begin{aligned} y_1 &= x^{\lambda_1} l_{11}(x), & y_2 &= x^{\lambda_2} l_{12}(x), \\ z_1 &= x\frac{dy_1}{dx} = x^{\lambda_1} l_{21}(x), & z_2 &= x\frac{dy_2}{dx} = x^{\lambda_2} l_{22}(x), \end{aligned}$$

it is clear that the functions $l_{ij}(x)$ are single-valued as well as analytic in the vicinity of $x = \infty$. To see this one has merely to observe that x^{λ_1} and x^{λ_2} are respectively multiplied by ρ_1 and ρ_2 when x makes a positive circuit of $x = \infty$. Furthermore the expression

$$l_{11}(x)\, l_{22}(x) - l_{12}(x)\, l_{21}(x) = x^{1-\lambda_1-\lambda_2}(y_1 y_2' - y_2 y_1') = x^{1-\lambda_1-\lambda_2} e^{-\int p dx}$$

is not zero for $|x| \geqq R$, since $p(x)$ is analytic. Consequently these functions $l_{11}(x), l_{12}(x), l_{21}(x), l_{22}(x)$ satisfy all the conditions of the auxiliary theorem.

Now let $a_{11}(x), a_{12}(x), a_{21}(x), a_{22}(x), e_{11}(x), e_{12}(x), e_{21}(x), e_{22}(x)$ be the functions and k_1, k_2 the integers fulfilling the fundamental equations (3) of the auxiliary theorem.

The pair of functions $l_{11}(x)$ and $l_{21}(x)$ depend on the initial choice of λ_1. If we multiply the equations (3) whose left-hand members are $l_{11}(x)$ and $l_{21}(x)$ by x^{-k_1}, these functions are replaced by $x^{-k_1} l_{11}(x)$ and $x^{-k_1} l_{21}(x)$ respectively. But this pair of functions is the same as the pair $l_{11}(x)$ and $l_{21}(x)$ that would be obtained if λ_1 were replaced by $\lambda_1 + k_1$. A similar remark applies to the pair of functions $l_{12}(x)$ and $l_{22}(x)$. That is, by properly choosing the four functions $l_{11}(x), l_{12}(x), l_{21}(x), l_{22}(x)$ we may take k_1 and k_2 equal to zero, and obtain

$$(11) \qquad \begin{aligned} l_{11}(x) &= a_{11}(x)\, e_{11}(x) + a_{12}(x)\, e_{21}(x), \\ l_{12}(x) &= a_{11}(x)\, e_{12}(x) + a_{12}(x)\, e_{22}(x), \\ l_{21}(x) &= a_{21}(x)\, e_{11}(x) + a_{22}(x)\, e_{21}(x), \\ l_{22}(x) &= a_{21}(x)\, e_{12}(x) + a_{22}(x)\, e_{22}(x), \end{aligned}$$

where the functions $a_{ij}(x)$ and $e_{ij}(x)$ have the properties stated in the auxiliary theorem.

Let us make this choice of the functions $a_{11}(x)$, $a_{12}(x)$, $a_{21}(x)$, $a_{22}(x)$ in the transformation (5). Multiplying the equations (11) for $l_{11}(x)$, $l_{21}(x)$ and $l_{12}(x)$, $l_{22}(x)$ by x^{λ_1} and x^{λ_2} respectively, we obtain

$$\begin{aligned} y_i &= a_{11}(x)\, Y_i + a_{12}(x)\, Z_i \\ z_i &= a_{21}(x)\, Y_i + a_{22}(x)\, Z_i \end{aligned} \qquad (i = 1, 2), \tag{12}$$

where

$$Y_1 = x^{\lambda_1} e_{11}(x), \quad Z_1 = x^{\lambda_1} e_{21}(x), \quad Y_2 = x^{\lambda_2} e_{12}(x), \quad Z_2 = x^{\lambda_2} e_{22}(x). \tag{13}$$

Hence, if the transformation (5) be chosen in this way, (Y_1, Z_1) and (Y_2, Z_2) are a pair of linearly independent solutions of the transformed system (6) of the form (13) where $e_{11}(x)$, $e_{12}(x)$, $e_{21}(x)$, $e_{22}(x)$ are entire functions such that $e_{11}(x)\, e_{22}(x) - e_{12}(x)\, e_{21}(x)$ is nowhere zero in the finite plane.

It is now an easy matter to determine explicitly the form of the coefficients $P_{ij}(x)$ in (6). If we substitute the above forms (13) for the known solutions in the first equation (6) we obtain at once

$$\frac{de_{11}(x)}{dx} + \frac{\lambda_1}{x} e_{11}(x) = P_{11}(x)\, e_{11}(x) + P_{12}(x)\, e_{21}(x),$$

$$\frac{de_{12}(x)}{dx} + \frac{\lambda_2}{x} e_{12}(x) = P_{11}(x)\, e_{12}(x) + P_{12}(x)\, e_{22}(x).$$

On account of the fact that the functions $e_{ij}(x)$ are entire, and that $e_{11}(x)\, e_{22}(x) - e_{12}(x)\, e_{21}(x)$ is nowhere zero, it follows from these equations that $P_{11}(x)$ and $P_{12}(x)$ are single-valued and analytic everywhere in the finite plane except for a possible pole of the first order at $x = 0$; by using the second equation (6) we may likewise show that $P_{21}(x)$ and $P_{22}(x)$ have these same properties. But the functions $P_{ij}(x)$ are analytic at $x = \infty$ according to the formulas (8), which hold inasmuch as $a_{11}(x)$, $a_{12}(x)$, $a_{21}(x)$, $a_{22}(x)$ are analytic at $x = \infty$ and reduce to $1, 0, 0, 1$ respectively at $x = \infty$. Accordingly the functions $P_{ij}(x)$ are linear functions of $1/x$, and are given by the series (8) which break off at the term in $1/x$. Therefore the equations (6) have the simple form

$$x \frac{dY}{dx} = rZ, \qquad x \frac{dZ}{dx} = sY + (x + 1 - p_1)\, Z, \tag{14}$$

where r and s are certain constants to be studied later. These equations can be integrated by a quadrature in the two cases $r = 0$ and $s = 0$.

Bearing in mind the nature of the transformation (5) that takes the original equations (4) into (14), and also that rZ and z may be replaced by

$x\,(dY/dx)$ and $x\,(dy/dx)$ respectively, we are led to summarize our result as follows:

THEOREM I. *If y satisfies the equation*

$$\frac{d^2y}{dx^2}+p(x)\frac{dy}{dx}+q(x)\,y=0,$$

$$p(x)=-1+\frac{p_1}{x}+\cdots, \qquad q(x)=\frac{q_2}{x^2}+\frac{q_3}{x^3}+\cdots,$$

then y and $x\,(dy/dx)$ can be represented in the forms

$$(15)\quad y=a_{11}(x)\,Y+a_{12}(x)\frac{x}{r}\frac{dY}{dx}, \qquad x\frac{dy}{dx}=a_{21}(x)\,Y+a_{22}(x)\frac{x}{r}\frac{dY}{dx}$$

where Y is some solution of

$$(16)\qquad \frac{d^2Y}{dx^2}+\left(-1+\frac{p_1}{x}\right)\frac{dY}{dx}-\frac{rs}{x^2}Y=0 \qquad (r, s \text{ constants}),$$

and $a_{11}(x)$, $a_{12}(x)$, $a_{21}(x)$, $a_{22}(x)$ are analytic at $x=\infty$, reducing to 1, 0, 0, 1 respectively at $x=\infty$. In the case $r=0$ it is necessary to replace

$$\frac{x}{r}\frac{dY}{dx} \text{ by } \lim_{r=0}\frac{x}{r}\frac{dY}{dx}.$$

The equation (16) is obtained from (14) under the hypothesis $r\neq 0$. When r approaches zero, it is possible to make the solution (Y, Z) of (14) approach any particular solution of that equation for $r=0$, i. e., one may write

$$Z=\lim_{r=0}\frac{x}{r}\frac{dY}{dx}.$$

A slight modification in the proof must be made in case the second alternative (9) is before us. In this event we may define $l_{11}(x)$, $l_{12}(x)$, $l_{21}(x)$, $l_{22}(x)$ so as to satisfy the restrictions of the auxiliary theorem by writing instead of (10)

$$y_1=x^{\lambda_1}l_{11}(x), \qquad y_2=x^{\lambda_1}\left(\frac{l_{11}(x)}{2\pi\rho_1\sqrt{-1}}\log x+l_{12}(x)\right),$$

$$z_1=x\frac{dy_1}{dx}=x^{\lambda_1}l_{21}(x), \qquad z_2=x\frac{dy_2}{dx}=x^{\lambda_1}\left(\frac{l_{21}(x)}{2\pi\rho_1\sqrt{-1}}\log x+l_{22}(x)\right).$$

From this point on we may develop the argument as we did before, except for a mere difference of analytical detail.

The equation (16), which comes from the linear system (14) by an elimination of Z, may be termed the *canonical form of equation* (1).

It is clear that a study of the solutions of (1) in the vicinity of $x=\infty$

reduces at once to a study of the solutions of this simple equation (16) which involves only the known constant p_1 and the product rs. The point $x = 0$ is a regular singular point of (16) and of (14) with exponents λ_1 and λ_2 by (13). The indicial equation at $x = 0$ is

$$\lambda^2 + (p_1 - 1)\lambda - rs = 0,$$

so that we have the relations

$$\lambda_1 + \lambda_2 = 1 - p_1, \qquad \lambda_1 \lambda_2 = -rs. \tag{17}$$

One may easily verify that these relations continue to hold even in the second alternative (9) when the multipliers are equal.

§ 2. *The Formal Solutions and the Exponents.*

The equation (1) admits of two *formal solutions** of the following types

$$s_1(x) = \left(1 + \frac{s_1^{(1)}}{x} + \frac{s_1^{(2)}}{x^2} + \cdots\right), \qquad s_2(x) = e^x x^{-p_1}\left(1 + \frac{s_2^{(1)}}{x} + \frac{s_2^{(2)}}{x^2} + \cdots\right).$$

This fact may be verified by a direct substitution which permits one to determine $s_1^{(1)}, s_{21}^{(2)}, \cdots$ and $s_2^{(1)}, s_2^{(2)}, \cdots$ uniquely so as to make $s_1(x)$ and $s_2(x)$ formally satisfy (1). Moreover the canonical equation (16) is also of this type (1), and in this case an explicit determination of these formal series $S_1(x)$, $S_2(x)$ gives at once†

$$S_1(x) = \left(1 - \frac{\lambda_1\lambda_2}{1}\cdot\frac{1}{x} + \frac{\lambda_1(\lambda_1+1)\lambda_2(\lambda_2+1)}{1\cdot 2}\cdot\frac{1}{x^2} + \cdots\right),$$

$$S_2(x) = e^x x^{-p_1}\left(1 + \frac{\lambda_1 - 1)(\lambda_2 - 1)}{1}\frac{1}{x} + \frac{(\lambda_1 - 1)(\lambda_1 - 2)(\lambda_2 - 1)(\lambda_2 - 2)}{1\cdot 2}\frac{1}{x^2} + \cdots\right). \tag{18}$$

Now, putting aside for the moment the case $r = 0$, we see that the formal solutions of the linear differential system (4) are given by

$$y = s_1(x), \qquad z = x\frac{ds_1(x)}{dx}; \qquad y = s_2(x), \qquad z = x\frac{ds_2(x)}{dx},$$

* A discussion of such formal solutions is given by HORN: *Gewöhnliche Differentialgleichungen beliebiger Ordnung*, pp. 183–188.

† An easy way to obtain the second series is to substitute $y = e^x x^{-p_1} w$ in (6), and in the resulting transformed equation put $w = \sum_{\nu=0}^{\infty} s_2^{(\nu)} x^{-\nu}$. The simplification of the form of the recurrent relation between the successive coefficients in both series depends on (17).

and that these solutions for (14) are

$$Y = S_1(x), \qquad Z = \frac{x}{r}\frac{dS_1(x)}{dx}; \qquad Y = S_2(x), \qquad Z = \frac{x}{r}\frac{dS_2(x)}{dx}.$$

It is also clear that each of these sets for Y, Z goes into the corresponding set for y, z, by the transformation (5). In fact the formal solutions Y, Z go into formal solutions y, z of the proper form with the same leading terms since the functions $a_{11}(x)$, $a_{12}(x)$, $a_{21}(x)$, $a_{22}(x)$ are power series in negative powers of x with leading coefficients 1, 0, 0, 1 respectively. It is directly demonstrable that the form and leading terms characterize the formal series for y, z completely, and hence the correspondence must be as stated.

Therefore if we replace y by $s_i(x)$ in (15) and Y by $S_i(x)$ respectively for $i = 1, 2$ the resulting equations are formally correct. *These relations permit us to determine power series for* $a_{11}(x)$, $a_{12}(x)/r$, $a_{21}(x)/r$, $a_{22}(x)/r$ *by means of the known series for* $s_i(x)$ *to be obtained from* (1), *and the series* $S_i(x)$ *given in* (18).

These latter series depends on λ_1 and λ_2 alone, so that we are in a position to actually obtain the functions $a_{11}(x)$, $a_{12}(x)/r$, $a_{21}(x)$, $a_{22}(x)/r$ in (15) as soon as λ_1 and λ_2 are known. That is, once a set of exponents λ_1, λ_2 are obtained it will be possible to determine the transformation (5) explicitly. We observe that, since $a_{22}(x)$ reduces to 1 at $x = \infty$, the power series for $a_{22}(x)/r$ begins with the term $1/r$, which permits us to determine r; and that when r is found, s may be determined by means of the relation $rs = -\lambda_1\lambda_2$.

Likewise even if r is zero the equations (14) again admit of similar formal solutions, and by substituting in both sides of (5) corresponding formal solutions of (4) and (14) one may determine $a_{11}(x)$, $a_{12}(x)$, $a_{21}(x)$, $a_{22}(x)$, and s.

It is essential therefore to determine the possible values of λ_1 and λ_2, and to know in what cases r may be zero. Concerning these exponents we know at present that

$$(19)\quad \lambda_1 = -\frac{1}{2\pi\sqrt{-1}}\log\rho_1, \qquad \lambda_2 = -\frac{1}{2\pi\sqrt{-1}}\log\rho_2, \qquad \lambda_1 + \lambda_2 = 1 - p_1.$$

Any pair of exponents λ_1 and λ_2 may be replaced by $\lambda_1 + m$ and $\lambda_2 - m$ respectively where m is any integer, without destroying the relations (19). The multipliers ρ_1, ρ_2 may be computed in various ways and we shall suppose these to be given quantities.

In order to completely consider all possibilities we need to make a subdivision into two cases, the first of which is the following: *If neither of the multipliers* ρ_1, ρ_2 *reduce to* 1 (Case *A*) *the exponents* λ_1, λ_2 *are unrestricted save*

as indicated by (19) *and neither r nor s is zero. In this case both formal solutions of* (1) *diverge.*

To prove this we note that there is a transformation of the form (5)

$$y = b_{11}(x)\,\bar{Y} + b_{12}(x)\,\bar{Z}, \qquad Z = b_{21}(c)\,\bar{Y} + b_{22}(x)\,\bar{Z}, \tag{20}$$

where $b_{11}(x)$, $b_{12}(x)$, $b_{21}(x)$, $b_{22}(x)$ are analytic at $x = \infty$ and reduce to $1, 0, 0, 1$ respectively at $x = \infty$, that takes the equations (14) into like equations in which the exponents λ_1, λ_2 are changed to $\lambda_1 - 1$, $\lambda_2 + 1$ respectively. In fact if we write

$$\begin{aligned} Y &= \left(1 + \frac{\lambda_1 - \lambda_2 - 1}{x}\right)\bar{Y} + \frac{\lambda_1}{s}\left(\frac{\lambda_1 - \lambda_2 - 1}{x}\right)\bar{Z}, \\ Z &= \frac{-s}{\lambda_1}\left(\frac{\lambda_1 - \lambda_2 - 1}{x}\right)\bar{Y} + \left(1 - \frac{\lambda_1 - \lambda_2 - 1}{x}\right)\bar{Z}, \end{aligned} \tag{21}$$

these equations (14) go into

$$x\frac{d\bar{Y}}{dx} = \frac{\lambda_1(1-\lambda_1)}{s}\bar{Z}, \qquad x\frac{d\bar{Z}}{dx} = \frac{s(1+\lambda_2)}{\lambda_1}\bar{Y} + (x + 1 - p_1)\bar{Z}$$

in which $-rs = \lambda_1\lambda_2$ is replaced by $-\overline{rs} = (\lambda_1 - 1)(\lambda_2 + 1)$. It is to be observed that r and s are not zero, for that implies that either λ_1 or λ_2 is zero by (17) so that one of the multipliers is 1.

By compounding the transformation (5) with the transformation (21) we obtain another transformation of the same form as (5) taking (4) directly to (21). Hence we have increased one of the exponents by unity and decreased the other by unity. By a succession of such steps we can obtain an equation (14) in which the exponents λ_1, λ_2 are replaced by $\lambda_1 + m$, $\lambda_2 - m$ where m is any integer. This proves our first statement above.

Now the formal series (18) diverge unless they terminate, and they cannot terminate since neither λ_1 nor λ_2 is an integer, by hypothesis. But it is obvious that corresponding series s and S for (1) and (16) either converge or diverge together by (5). It follows that the two formal solutions of (1) diverge in this case.

The second case is the following: *If one of the multipliers, as* ρ_1, *reduces to* 1 (Case *B*) *it is possible to take* $\lambda_1 = 0$ *and* $\lambda_2 = 1 - p_1$. *At least one of the formal solutions* $s_1(x)$, $s_2(x)$ *of* (1) *converges. If* $s_1(x)$ *diverges, then r but not s reduces to zero; if* $s_2(x)$ *diverges, s but not r reduces to zero; if both* $s_1(x)$ *and* $s_2(x)$ *converge, we may take* $r = s = 0$.

In this case one of the exponents is zero at the outset, or by a succession of transformations (20) it will be possible to reduce an integral exponent corresponding to the multiplier 1 to the value zero so that $\lambda_1 = 0$; we shall have then $\lambda_2 = 1 - p_1$, by (17).

Also since we have $rs = -\lambda_1\lambda_2 = 0$, either r or s is zero.

If $r = 0$, the system (14) becomes

$$x\frac{dY}{dx} = 0, \qquad x\frac{dZ}{dx} = sY + (x + 1 - p_1)\,Z. \tag{22}$$

The two formal solutions of this system are clearly

$$\begin{aligned} &Y = 1, \qquad Z = -\frac{s}{x}\left(1 + \frac{p_1 - 2}{x} + \frac{(p_1 - 2)(p_1 - 3)}{x^2} + \cdots\right), \\ &Y = 0, \qquad Z = e^x x^{1-p_1}, \end{aligned} \tag{23}$$

the second of which is terminating and so convergent. If $s = 0$ the system (14) has the two formal solutions

$$\begin{aligned} &Y = 1, \qquad Z = 0, \\ &Y = re^x x^{-p_1}\left[1 + \frac{p_1}{x} + \frac{p_1(p_1 + 1)}{x^2} + \cdots\right], \qquad Z = e^x x^{1-p_1}, \end{aligned} \tag{24}$$

by (18), and here the first series is terminating and convergent. In either case one of the formal solutions of (14), and hence the corresponding formal solution of (1) converges. If $s_1(x)$ diverges, so will the corresponding solution of (16) and we must have the first case $r = 0$. Since s appears as a factor in the first series (23) we infer $s \neq 0$ in this case. Likewise if $s_2(x)$ diverges we infer $r \neq 0$, $s = 0$.

It only remains therefore to show that in the case when both $s_1(x)$ and $s_2(x)$ converge we may take $r = s = 0$. We have already seen that either r or s is zero. Suppose we have $r = 0, s \neq 0$. An inspection of the first formal solution (23) shows that p_1 must have one of the values $2, 3, \cdots$ in order that the series for Z may terminate and converge. The first solution may be written $Y = 1, Z = -sP(x)/x$, where $P(x)$ is a polynomial in $1/x$. If we make the further transformation

$$Y = \bar{Y}, \qquad Z = -\frac{s}{x}P(x)\,\bar{Y} + \bar{Z},$$

of the form (20) where $b_{11}(x), b_{12}(x), b_{21}(x), b_{22}(x)$ are analytic at $x = 0$ and reduce to $1, 0, 0, 1$ respectively at $x = \infty$, the resulting system in $\bar{Y}, \bar{Z}$ is

$$x\frac{d\bar{Y}}{dx} = 0, \qquad x\frac{d\bar{Z}}{dx} = (x + 1 - p_1)\,\bar{Z}$$

of the form (14) with $r = s = 0$. But the transformation (5) and the above transformation are together equivalent to a single transformation of the form (5), so that we may take $r = s = 0$. Likewise we may dispose of the case $r \neq 0, s = 0$ by a consideration of the second of the solutions (24).

§ 3. *Explicit Solution of the Canonical Form of Equation.*

In order to study completely the solutions of (1) in the vicinity of $x = \infty$ it is necessary to solve the canonical form (16) of equation equivalent to the linear system (14). If either multiplier ρ_1, ρ_2 is 1, we have essentially Case *B* and we may reduce the equation (14) to a form in which $r = 0$ or $s = 0$. In either case the reduced system (14) is integrable by a quadrature, and we have

$$(25)\qquad \begin{aligned} &Y = c, \qquad Z = e^x x^{1-p_1}\Big[d + cs\int e^{-x} x^{p_1-2}\,dx\Big] && \text{if } r = 0,\\ &Y = d + cr\int e^x x^{-p_1}\,dx, \qquad Z = ce^x x^{1-p_1} && \text{if } s = 0. \end{aligned}$$

Therefore the canonical equations (14) can always be solved by a single quadrature if either λ_1 or λ_2 is an integer.

However when neither λ_1 nor λ_2 is an integer it is possible to effect the solution of (16) by means of the Laplace transformation. To simplify the form of this transformation let us write $Y = x^{\lambda_1}\bar{Y}$ in (12) and obtain the equation*

$$(26)\qquad x\frac{d^2\bar{Y}}{dx^2} + (1 + \lambda_1 - \lambda_2 - x)\frac{d\bar{Y}}{dx} - \lambda_1\bar{Y} = 0.$$

This equation is of the simple type to which Laplace first applied his transformation, and using his method we find the formulas

$$(27)\qquad Y_i = x^{\lambda_1}\int_{l_i} e^{xz} z^{\lambda_1-1}(1-z)^{-\lambda_2}\,dz \qquad (i = 1, 2),$$

and likewise by interchanging λ_1 and λ_2, we find

$$(28)\qquad Y_i = x^{\lambda_2}\int_{l_i} e^{xz} z^{\lambda_2-1}(1-z)^{-\lambda_1}\,dz \qquad (i = 1, 2).$$

Here the contours l_1 and l_2 are loop-circuits to the points $z = 0$ and $z = 1$ respectively taken from $z = \infty$ along a ray to these points, making a circuit in a negative sense and returning to ∞ along the ray. The ray must be so chosen that the integrals converge.

Under the conditions that neither λ_1 nor λ_2 is an integer the two solutions of either set are linearly independent.

* Horn has considered the application of the Laplace transformation to the equation

$$(a_0 + b_0 x)\frac{d^2y}{dx} + (a_1 + b_1 x)\frac{dy}{ax} + (a_2 + b_2 x)\,y = 0,$$

which he reduces first to a normal form somewhat different from our equation (26): Mathematische Annalen, vol. 49 (1907), pp. 453–492. The reader who desires to see how the integral given by the Laplace transformation may be made to yield complete information about the solution Y is referred to this interesting paper.

It is now possible to give more explicit form to the relation (15) of Theorem I which we do in the following:

THEOREM I'. *Two linearly independent solutions of an equation*

$$\frac{d^2y}{dx^2} + p(x)\frac{dy}{dx} + q(x)\,y = 0,$$

$$p(x) = -1 + \frac{p_1}{x} + \cdots, \qquad q(x) = \frac{q_2}{x^2} + \frac{q_3}{x^3} + \cdots$$

may be represented in the form

$$y_i = x^{\lambda_1}\left[a_i(x)\int_{l_i} e^{xz}\,z^{\lambda_1-1}(1-z)^{-\lambda_2}\,dz + b_i(x)\int_{l_i} e^{xz}\,z^{\lambda_1}(1-z)^{-\lambda_2}\,dz\right] \qquad (i = 1, 2),$$

where $a_i(x)$ and $b_i(x)$ are analytic at $x = \infty$ and reduce to 1 and 0 respectively at $x = \infty$, and λ_1, λ_2 are any determinations of λ_1, λ_2 satisfying the equations (18). *This representation breaks down only when one of the multipliers ρ_1, ρ_2 reduces to 1, and then it may be replaced by one of the two following:*

$$y_1 = a(x) + b(x)\,e^x\,x^{1-p_1}\int e^{-x}\,x^{p_1-2}\,dx, \qquad y_2 = b(x)\,e^x\,x^{1-p_1},$$

$$y_1 = a(x)\int e^x\,x^{-p_1}\,dx + b(x)\,e^x\,x^{1-p_1}, \qquad y_2 = a(x).$$

To prove these relations it is only necessary to write in Theorem I $a_{11}(x) = a(x)$, $a_{12}(x)/r = b(x)$ and to use for Y and Z the forms deduced from (28) and (25).

§ 4. *Characterization of the Laurent Series for the Solutions.*

Let us now restrict ourselves to the non-exceptional case when both of the multipliers ρ_1, ρ_2 are distinct from each other and from 1. In this case the equation (1) has a pair of solutions $y_1(x), y_2(x)$ of the forms

$$y_1(x) = x^{\lambda_1}\,l_{11}(x), \qquad y_2(x) = x^{\lambda_2}\,l_{12}(x),$$

where $l_{11}(x)$ and $l_{12}(x)$ may be expressed as Laurent series

$$l_{11}(x) = \sum_{\nu=-\infty}^{+\infty} l_1^{(\nu)}\,x^\nu, \qquad l_{12}(x) = \sum_{\nu=-\infty}^{+\infty} l_2^{(\nu)}\,x^\nu,$$

since these functions are single-valued and analytic for $|x| \geqq R$.

It is an interesting matter to determine the precise character of the coefficients $l_i^{(\nu)}$ in these series. To do so we make use of the relations obtained from the transformation (5) taking the linear system (4) to the canonical system (14). The first equation of transformation may be written

$$y = a(x)\,Y + xb(x)\frac{dY}{dx}, \tag{29}$$

where Y satisfies (14) and $a(x)$ and $b(x)$ are analytic at $x = \infty$ and reduce to 1 and 0 respectively at $x = \infty$. Let us express $a(x)$ and $b(x)$ in the form of power series in negative powers of x

$$a(x) = 1 + \sum_{\nu=1}^{\infty} a_\nu x^{-\nu}, \qquad b(x) = \sum_{\nu=1}^{\infty} b_\nu x^{-\nu}. \tag{30}$$

It is clear that the special solutions Y_1 and Y_2 of (16) which correspond to the exponents λ_1 and λ_2 go over into the solutions y_1 and y_2 respectively, so that we may write

$$y_i = a(x)\,Y_i + xb(x)\frac{dY_i}{dx} \qquad (i = 1, 2).$$

But the functions $x^{-\lambda_1}Y_1$, $x^{-\lambda_2}Y_2$ admit of expansion in Laurent series in x convergent for all finite values of x, since (16) has no finite singular points except the regular point $x = 0$ with exponents λ_1, λ_2.

If we write $Y = x^{\lambda_i}\bar{Y}$ and substitute

$$\bar{Y}_i = \sum_{\nu=0}^{\infty} L_i^{(\nu)} x^\nu \qquad (i = 1, 2),$$

in (26) we find that

$$\begin{aligned} Y_1 &= x^{\lambda_1}\left(1 + \frac{\lambda_1}{1\cdot(\lambda_1-\lambda_2+1)}x + \frac{\lambda_1\cdot(\lambda_1+1)}{1\cdot 2\cdot(\lambda_1-\lambda_2+1)\cdot(\lambda_1-\lambda_2+2)}x^2 + \cdots\right), \\ Y_2 &= x^{\lambda_2}\left(1 + \frac{\lambda_2}{1\cdot(\lambda_2-\lambda_1+1)} + \frac{\lambda_2\cdot(\lambda_2+1)}{1\cdot 2\cdot(\lambda_2-\lambda_1+1)\cdot(\lambda_2-\lambda_1+2)}x^2 + \cdots\right). \end{aligned} \tag{30}$$

Putting in the explicit series for y_i, $a(x)$, $b(x)$, Y_i in (29) and comparing coefficients we obtain the following:

THEOREM II. *If the multipliers ρ_1, ρ_2 are distinct from each other and from 1, and if λ_1, λ_2 are any pair of exponents satisfying the conditions* (19), *the coefficients $l_i^{(\nu)}$ in the Laurent series*

$$y_1 = x^{\lambda_1}\sum_{\nu=-\infty}^{+\infty} l_1^{(\nu)} x^\nu, \qquad y_2 = x^{\lambda_2}\sum_{\nu=-\infty}^{+\infty} l_2^{(\nu)} x^\nu,$$

in which two particular linearly independent solutions y_1, y_2 of (1) *may be expanded, have the form*

$$l_i^{(\nu)} = L_i^{(\nu)} + [a_1 + b_1(\lambda_i + \nu + 1)]\,L_i^{(\nu+1)} + [a_2 + b_2(\lambda_i + \nu + 2)]\,L_i^{(\nu+2)} + \cdots,$$

where

$$L_1^{(\nu)} = \frac{\lambda_1 \cdot (\lambda_1 + 1) \cdots (\lambda_1 + \nu - 1)}{1 \cdot 2 \cdots \nu \cdot (\lambda_1 - \lambda_2 + 1) \cdot (\lambda_1 - \lambda_2 + 2) \cdots (\lambda_1 - \lambda_2 + \nu)}$$

(32)

$$L_2^{(\nu)} = \frac{\lambda_2 \cdot (\lambda_2 + 1) \cdots (\lambda_2 + \nu - 1)}{1 \cdot 2 \cdots \nu \cdot (\lambda_2 - \lambda_1 + 1) \cdot (\lambda_2 - \lambda_1 + 2) \cdots (\lambda_2 - \lambda_\nu + \nu)}$$

and a_ν, b_ν *are a set of numbers such that* $\sqrt[\nu]{a_\nu}$ *and* $\sqrt[\nu]{b_\nu}$ *are limited in absolute value for* $\nu = 1, 2, \cdots$.

The condition on a_ν and b_ν stated in this theorem is equivalent to saying that the series for $a(x)$ and $b(x)$ converge for $|x|$ sufficiently large, according to a well-known test due to Cauchy. The form for $l_i^{(\nu)}$ given in the theorem is sufficient to describe these coefficients completely.

It is not difficult to show that a similar characterization may be made in the special cases not coming under the hypothesis of the theorem.

§ 5. *Functional Relations.*

The theorems which have been developed in the preceding paragraphs are based on a preliminary normalizing transformation of the dependent variable in (1). This disposes of all but two essentially arbitrary constants. The transformation of *independent* variable $x = \bar{x} + \psi(\bar{x})$, where $\psi(\bar{x})$ is analytic at $\bar{x} = \infty$, leaves (1) unaltered in form but leads to no additional normalization. This fact is expressed in the following:

THEOREM III. *If the multipliers* ρ_1, ρ_2 *are distinct from each other and from* 1, *and if the function* $\varphi(x)$ *has the form* $x + \psi(x)$ *where* $\psi(x)$ *is analytic at* $x = \infty$, *then we have for every solution* $Y(x)$ *of* (16)

$$Y(\varphi(x)) = a(x)\, Y(x) + b(x) \frac{dY(x)}{dx} \tag{33}$$

where $a(x)$ *and* $b(x)$ *are analytic at* $x = \infty$.

The proof is immediate. Such a transformation applied to (16) takes that equation into an equation of the form (1) for which ρ_1, ρ_2 are the same as for (16) and whose formal solutions clearly diverge (Case *A*) with the corresponding formal solutions of (16), inasmuch as the new formal solutions result from the original formal solutions by this same transformation $x = \varphi(\bar{x})$. Consequently according to the results of § 2 it is possible to take the canonical equation associated with this transformed equation as in Theorem I to be the same as the equation (16) in Y itself.

If now we replace y, $a_{11}(x)$, and $xa_{12}(x)/r$ by $Y(\varphi(x))$, $a(x)$, and $b(x)$ respectively in (15) we obtain the equation (33) of the above theorem, where $a(x)$ reduces to 1 at $x = \infty$. Moreover the Y on both sides of equation (33) refers always to one and the same solution, except for a constant

multiplier; this is a consequence of the fact that $Y(\varphi(x))$ and $Y(x)$ undergo similar transformations when x makes a positive circuit of $x = \infty$. Such a constant multiplier may be absorbed nto $a(x)$ and $b(x)$ so that the modified equation (33) has the desired form.

It is manifest that simple analogues of Theorems II and III exist in the exceptional case when the multipliers become equal or reduce to unity

ON QUADRATIC RESIDUES*

BY

J. McDONNELL

1. Let m be any integer, n a positive odd integer, and s a primitive nth root of unity. We shall prove that the product

$$A = \prod_{k=1}^{\frac{1}{2}(n-1)} \frac{s^{km} - s^{-km}}{s^k - s^{-k}}$$

is identical with Jacobi's symbol (m/n) and obtain a proof of the quadratic reciprocity theorem. If m and n have a common factor g, then $A = 0$, since the numerator of A has the factor $s^{qm} - s^{-qm} = 0$, where $q = n/g$.

2. We shall show that, when m and n are relatively prime, A is independent of the particular primitive root s employed. First, let m be positive and odd, and let r be a primitive mth root of unity. Since m is odd, the mth roots of unity are $1, r^2, r^4, \cdots, r^{2(m-1)}$. Hence

$$x^m - y^m \equiv (x - y)(xr - yr^{-1})(xr^2 - yr^{-2}) \cdots (xr^{m-1} - yr^{-(m-1)}),$$

identically. Take $x = s^q$, $y = s^{-q}$. We see that

$$A = \prod_{p=1}^{m-1} \prod_{q=1}^{\frac{1}{2}(n-1)} (r^p s^q - r^{-p} s^{-q}).$$

Similarly,

$$B = \prod_{k=1}^{\frac{1}{2}(m-1)} \frac{r^{kn} - r^{-kn}}{r^k - r^{-k}} = \prod_{p'=1}^{\frac{1}{2}(m-1)} \prod_{q'=1}^{n-1} (r^{p'} s^{q'} - r^{-p'} s^{-q'}).$$

To any factor $F = r^p s^q - r^{-p} s^{-q}$ in A there corresponds a factor

$$F' = r^{p'} s^{q'} - r^{-p'} s^{-q'}$$

in B such that one of the alternatives

(i) $\quad p = p', \quad q = q', \quad p \leqq \frac{1}{2}(m-1), \quad q' \leqq \frac{1}{2}(n-1),$

(ii) $\quad p + p' = m, \quad q + q' = n, \quad p > \frac{1}{2}(m-1), \quad q' > \frac{1}{2}(n-1),$

holds. In the first case, $F = F'$; in the second, $F = -F'$. Hence

$$A = (-1)^{\frac{(m-1)(n-1)}{4}} B.$$

As B is independent of s, A is also.

* Presented to the Society (Chicago), March 21, 1913.

Next, let m be negative or even. Choose t so that $\mu = m + tn$ is positive and odd. The expression in § 1 for A remains unaltered if we replace s^m by s^μ. By the preceeding proof, A is independent of s.

Accordingly we can in all cases represent the function A by the symbol $f(m, n)$, since it depends upon m and n only.

3. We shall now show how to evaluate $f(m, n)$ by a process of reduction based upon the following three properties:*

$$(1) \qquad f(m, n) = f(p, n) \qquad \text{if } p \equiv m \pmod{n},$$

$$(2) \qquad f(p, n) = (-1)^{\frac{n-1}{2}} f(q, n) \qquad \text{if } p + q \equiv 0 \pmod{n},$$

$$(3) \qquad f(m, n) = (-1)^{\frac{(m-1)(n-1)}{4}} f(n, m) \text{ if } m \text{ and } n \text{ are positive and odd.}$$

Let p be the positive remainder $< n$ obtained by dividing m by n. Then $f(m, n) = f(p, n)$ by (1). If p is even, define the odd number q by $p + q = n$ and apply (2). Thus either p itself or else q is positive and less than n. We now apply (3). After repetitions of this process, we ultimately reach unity and an odd number l such that $f(m, n) = \pm f(1, l)$. But $f(1, l) = 1$ (§ 1). Thus $f(m, n) = \pm 1$ and the sign is determined by the reduction process.

4. If $m = pq$, then

$$(4) \qquad f(m, n) = f(p, n) \cdot f(q, n).$$

We may assume that m is relatively prime to n, since otherwise each member is zero by § 1. Then $s_1 = s^q$ is a primitive nth root of unity and

$$\frac{s^{km} - s^{-km}}{s^k - s^{-k}} = \frac{s_1^{kp} - s_1^{-kp}}{s_1^k - s_1^{-k}} \cdot \frac{s^{kq} - s^{-kq}}{s^k - s^{-k}},$$

from which (4) follows. In particular,

$$(5) \qquad f(m, n) = \{f(2, n)\}^p \cdot f(q, n), \qquad \text{if } m = 2^p q \qquad (q \text{ odd}).$$

By (2) and (3),

$$f(2, n) = (-1)^{\frac{n-1}{2}} f(n-2, n), \qquad f(n-2, n) = f(n, n-2),$$

* The third property was proved in § 2. The first and second follow from

$$s^{kp} = s^{km}, \qquad s^{kq} - s^{-kq} = -s^{kp} + s^{-kp}.$$

since $(n-1)(n-3) \equiv 0 \pmod 8$. Then, from these and (1),

$$f(2,n) = (-1)^{\frac{n-1}{2}} f(2,n-2). \tag{6}$$

If ω is an imaginary cube root of unity,

$$f(2,3) = \frac{\omega^2 - \omega^{-2}}{\omega - \omega^{-1}} = -1.$$

Then, by (6), $f(2,5) = -1$, $f(2,7) = 1$, $f(2,9) = 1$, and, generally,

$$f(2,n) = (-1)^{\frac{n^2-1}{8}}. \tag{7}$$

5. If $n = ab$, as in any text on the theory of numbers,

$$\frac{n^2-1}{8} \equiv \frac{a^2-1}{8} + \frac{b^2-1}{8}, \qquad \frac{n-1}{2} \equiv \frac{a-1}{2} + \frac{b-1}{2} \pmod 2.$$

Hence by (7) and (3), with m replaced by an odd number q,

$$f(2,n) = f(2,a) \cdot f(2,b), \qquad f(q,n) = f(q,a) \cdot f(q,b).$$

The extension to the case in which n is the product $abc \cdots$ of several factors is evident. Hence, using (5), we get

$$f(m,n) = f(m,a) \cdot f(m,b) \cdot f(m,c) \cdots. \tag{8}$$

6. Let x be the product of the factors in the numerator of A (§ 1) and y the product of those in the denominator of A. Set $\eta = (-1)^{\frac{(n-1)(n-3)}{8}}$. As shown by GAUSS,*

$$\eta x = 1 + s^m + s^{3m} + \cdots + s^{m(n-1)^2}, \qquad \eta y = 1 + s + s^3 + \cdots + s^{(n-1)^2}.$$

First, let n be a prime, and let α range over the quadratic residues $(< n)$ of n, and β over the non-residues. Then

$$\eta y = 1 + 2\Sigma s^\alpha, \qquad \eta x = 1 + 2\Sigma s^\alpha \quad \text{or} \quad 1 + 2\Sigma s^\beta,$$

according as m is a quadratic residue or non-residue. Hence

$$x = y, \qquad f(m,n) = x/y = 1,$$

if m be a residue of n; while if m be a non-residue,

$$\eta(x+y) = 2(1 + \Sigma s^\alpha + \Sigma s^\beta) = 0, \qquad f(m,n) = -1.$$

Hence, if n is a prime, $f(m, n)$ is identical with Legendre's symbol (m / n).

Next, let $n = abc \cdots$, where $a, b, \cdots$ are primes. Then, by (8),

$$f(m, n) = \left(\frac{m}{a}\right)\left(\frac{m}{b}\right)\left(\frac{m}{c}\right)\cdots.$$

Thus $f(m, n)$ is identical with Jacobi's symbol (m / n).

* *Disquisitiones arithmeticae*, German translation by MASER, 1889, p. 474.

A SET OF FIVE INDEPENDENT POSTULATES FOR BOOLEAN ALGEBRAS, WITH APPLICATION TO LOGICAL CONSTANTS*

BY

HENRY MAURICE SHEFFER

Introduction.

Postulate-sets for determining the class of Boolean algebras† have been given by SCHRÖDER,‡ WHITEHEAD,§ and HUNTINGTON.|| SCHRÖDER'S set of ten postulates assumes—in addition to an undefined class K, common to all these postulate-sets—an undefined dyadic relation, $\nleq$, and BOOLE'S¶ undefined binary K-rules** of combination, $+$ and $\times$; WHITEHEAD'S two sets, the first of thirteen, and the second of fifteen, postulates, and HUNTINGTON'S first set, of ten postulates, assume the same undefined K-rules of combination, $+$ and $\times$, which HUNTINGTON writes respectively $\oplus$ and $\odot$; HUNTINGTON'S second set, of nine postulates, assumes SCHRÖDER'S undefined relation $\nleq$,

* Presented to the Society, December 31, 1912.

† We employ the term *Boolean algebras* in its plural form for the following reasons: (1) none of the equivalent postulate-sets here referred to is in terms of its undefined entities one-valued ("categorical")—that is, each determines not a single algebra but a class of algebras; one should not speak, therefore, of "*der* identische Kalkul" (SCHRÖDER), "*the* Algebra of Symbolic Logic" (WHITEHEAD), or "*the* algebra of logic" (HUNTINGTON); (2) PEANO'S *Formulario Mathematico* and WHITEHEAD and RUSSELL'S *Principia Mathematica*, each of which includes, as a part, the algebras under consideration, have a far stronger title to the name "algebra of logic"; (3) "The Algebra of Symbolic Logic, viewed as a distinct algebra, is due to Boole" (WHITEHEAD, *loc. cit.*, p. 115). "This algebra in all its essential particulars was invented and perfected by Boole" (*ib.*, p. 35, footnote).

‡ ERNST SCHRÖDER: *Vorlesungen über die Algebra der Logik (Exakte Logik)*, Erster Band, 1890. The postulates, under various names, are scattered throughout the volume; collected into one list by E. MÜLLER: *Abriss der Algebra der Logik*, Erster Teil (1909), pp. 20, 21.

§ A. N. WHITEHEAD: *A Treatise on Universal Algebra*, Vol. I, 1898, pp. 35–37.

|| E. V. HUNTINGTON: *Sets of Independent Postulates for the Algebra of Logic.* These Transactions, vol. 5 (1904), pp. 288–309.

¶ G. BOOLE: *An Investigation of the Laws of Thought.* London, 1854.

** An *n-ary rule of combination* ϕ, is an agreement according to which any n (distinct or non-distinct) logical entities, $a_1, a_2, a_3, \cdots$, in a definite order, determine a unique logical entity $\phi(a_1, a_2, a_3, \cdots)$; in other words, a rule of combination is a *one-valued* logical function. If the entity $\phi(a_1, a_2, a_3, \cdots)$ is defined for all those and only those cases where all the n entities $a_1, a_2, a_3, \cdots$ are elements of some class K, then ϕ is a *K-rule of combination*; for a *binary* K-rule of combination, $\phi(a, b)$ is also conveniently written $a \circ b$. If, for a K-rule of combination ϕ, the entity $\phi(a_1, a_2, a_3, \cdots)$ is always a K-element, then ϕ is *K-closed.*

which he writes $\circledcirc$; and his third set, of nine postulates, the undefined K-rule of combination $\oplus$. Each of these sets contains three existence-postulates, namely, those demanding the existence of (1) the special Boolean* element z; (2) the special Boolean* element u; and (3) for any K-element a, its corresponding Boolean element $\bar{a}$. The independence of all the postulates of each set is proved only for Huntington's sets; and Huntington was the first to show that any two of the concepts $\oplus$, $\odot$, and $\circledcirc$ are definable in terms of the third.

In this paper we offer, in § 1, a set of five independent postulates for Boolean algebras. This set, which like Huntington's third set assumes but one undefined K-rule of combination, differs from the previous sets (1) in the small number of postulates, and (2) in the fact that the set contains no existence-postulate for z, u, or $\bar{a}$.

In § 2 we apply our results to the problem of reducing the number of primitive logical constants.†

§ 1. *Postulate-Set for Boolean Algebras.*

We assume:

I. A class K,

II. A binary K-rule of combination $|$,

III. The following properties of K and $|$:

1. There are at least two distinct K-elements.
2. Whenever a and b are K-elements, $a \mid b$ is a K-element.

$$Def. \quad a' = a \mid a.$$

3. Whenever a and the indicated combinations of a are K-elements,

$$(a')' = a.$$

4. Whenever a, b, and the indicated combinations of a and b are K-elements,

$$a \mid (b \mid b') = a'.$$

5. Whenever a, b, c, and the indicated combinations of a, b, and c are K-elements,

$$(a \mid (b \mid c))' = (b' \mid a) \mid (c' \mid a).$$

For convenience, $a \mid b$ may be read *a per b*.

Classification of Postulates 1–5.

Postulate 1 is an existence-postulate. Postulate 2, which demands that the K-rule of combination $|$ shall be K-closed,‡ is a *K-closing* postulate.

* For z, BOOLE and SCHRÖDER write 0; for u, 1. *Cf.* p. 486, footnote †.

† WHITEHEAD and RUSSELL: *Principia Mathematica*, vol. I, 1910, pp. 94–101.

‡ See p. 481, footnote **.

Postulate 3, which demands that a and $(a')'$ shall always be names for the same K-element—that the names a and $(a')'$ shall always be equivalent—is an *equivalence* postulate; so are 4 and 5.

Thus our set consists of an existence-postulate, a K-closing postulate, and three equivalence postulates. Moreover, if we do not wish to exclude systems which have but a single element, then 1 may be replaced by the weaker postulate

1′. There is at least one K-element.

Consistence of Postulates 1–5.

With the following interpretation of K and $|$, postulates 1–5 are satisfied: K has only two distinct elements, m and n; $m|m=n$, $m|n=n|m=n|n=m$.

Independence of Postulates 1–5.

With each of the interpretations of K and $|$ given in (1)–(5) below, all the postulates, except the one correspondingly numbered, are satisfied; that postulate is, therefore, independent of the remaining four.

(1) K has only one element m; $m \mid m = m$.

(2) K has any number, greater than one, of distinct elements; for any K-element m, $m \mid m = m$; for any two distinct K-elements, m and n, $m \mid n$ is not a K-element.

(3) K has only two distinct elements, m and n; $m \mid m = m \mid n = n \mid m = n \mid n = m$.

(4) K is the class of all rationals; for any K-elements, m and n, $m \mid n = -\frac{1}{2}(m+n)$. Postulate 4 holds only when $m = 0$.

(5) K has only three distinct elements, l, m, and n; $|$ is defined by the following table (for example: $m \mid l = n$).

$\mid$	l	m	n
l	l	m	n
m	n	n	l
n	m	l	m

Deductions from Postulates 1–5.

The proofs are given after theorem V.

A. Whenever a and b are K-elements, $a \mid b = b \mid a$.

B. Whenever a and b are K-elements, $a \mid a' = b \mid b'$.

I*a*. Whenever a and b are K-elements, $(a \mid b)'$ is a K-element.

I*b*. Whenever a and b are K-elements, $a' \mid b'$ is a K-element.

II*a*. There is a K-element z such that for any K-element a, $(a \mid z)' = a$.

II*b*. There is a K-element u such that for any K-element a, $a' \mid u' = a$.

III*a*. Whenever $a, b, (a \mid b)'$, and $(b \mid a)'$ are K-elements, $(a \mid b)' = (b \mid a)'$.

III*b*. Whenever $a, b, a' \mid b'$, and $b' \mid a'$ are K-elements, $a' \mid b' = b' \mid a'$.

IV*a*. Whenever $a, b, c, (a \mid b)', (a \mid c)', b' \mid c', [a \mid (b' \mid c')]'$, and $[(a|b)'|'|[(a|c)'|']'$ are K-elements, $[a|(b'|c')]' = [(a|b)'|']'|[(a|c)'|']'$.

IV*b*. Whenever $a, b, c, a' \mid b', a' \mid c', (b \mid c)', a' \mid [(b \mid c)']'$, and $[(a'|b')|(a'|c')]'$ are K-elements, $a'|[(b|c)'|']' = [(a'|b')|(a'|c')]'$.

V*. If z and u of II*a* and II*b* are unique K-elements, then for any K-element a there is a K-element $\bar{a}$ such that $(a \mid \bar{a})' = u$ and $a' \mid (\bar{a})' = z$.

Proofs of the Preceding Theorems.

In the following proofs the use of postulate 2 is not always explicitly mentioned.

Proof of A.

$$
\begin{aligned}
a \mid b &= [(a \mid b)' \,]' && \text{[by 3]} \\
&= [\,(a \mid \{\, b' \,\}')' \,]' && \text{[by 3} \\
&= [\,(\{\, b' \,\}' \mid a)' \,]' && \text{[by 5, } b \text{ replaced by } b' \text{ and } c \text{ by } b'\text{]} \\
&= [\,(b \mid a)' \,]' && \text{[by 3]} \\
&= b \mid a && \text{[by 3].}
\end{aligned}
$$

Proof of B.

$$
\begin{aligned}
a \mid a' &= [\,(a \mid a')' \,]' && \text{[by 3]} \\
&= [\,(a \mid a') \mid (b \mid b') \,]' && \text{[by 4, } a \text{ replaced by } a \mid a'\text{]} \\
&= [\,(b \mid b') \mid (a \mid a') \,]' && \text{[by } A\text{]} \\
&= [\,(b \mid b')' \,]' && \text{[by 4, } a \text{ replaced by } b \mid b' \text{ and } b \text{ by } a\text{]} \\
&= b \mid b' && \text{[by 3].}
\end{aligned}
$$

Proof of Ia. Use 2 twice.

Proof of Ib. Use 2 thrice.

* Other theorems, the proofs of which we omit, are:

C. $(a \mid b) \mid (a \mid b') = a$.

D. $[a' \mid (a \mid b)]' = a \mid (a' \mid b')$.

E. $[(a \mid b') \mid (a' \mid b)]' = (a \mid b) \mid (a' \mid b')$.

F. $a \mid [a \mid (b \mid c)] = b \mid [b \mid (c \mid a)] = c \mid [c \mid (a \mid b)]$
$= a \mid (b \mid c)' \quad = b \mid (c \mid a)' \quad = c \mid (a \mid b)'$.

G. $(a' \mid a) \mid [(b' \mid a) \mid (c' \mid a)] = a\,(b \mid c)$.

Proof of IIa.

There is a K-element—say, x [by 1].

There is a K-element, $x \mid x'$, which we call z [by 2, used twice].

$$(a \mid z)' = (a')' \qquad \text{[by 4, } b \text{ replaced by } x\text{]}$$
$$= a\,. \qquad \text{[by 3].}$$

Proof of IIb.

There is a K-element, z', which we call u [by IIa and 2].

$$a' \mid u' = a' \mid (z')'$$
$$= a' \mid z \quad \text{[by 3]}$$
$$= (a')' \quad \text{[by 4, } a \text{ replaced by } a' \text{ and } b \text{ by } x\text{]}$$
$$= a \quad \text{[by 3].}$$

Proof of IIIa. $a \mid b = b \mid a$ [by A].

Hence $(a \mid b)' = (b \mid a)'$.

Proof of IIIb. $a' \mid b' = b' \mid a'$

[by A, a replaced by a' and b by b'].

Proof of IVa. $[\,a \mid (b' \mid c')\,]' = [\,(b')' \mid a\,] \mid [\,(c')' \mid a\,]$

[by 5, b replaced by b' and c by c']

$$= (b \mid a) \mid (c \mid a) \qquad \text{[by 3, used twice]}$$
$$= (a \mid b) \mid (a \mid c) \qquad \text{[by } A\text{, used twice]}$$
$$= [\,(a \mid b)'\,]' \mid [\,(a \mid c)'\,]' \qquad \text{[by 3, used twice].}$$

Proof of IVb.

$$a' \mid [\,(b \mid c)'\,]' = a' \mid (b \mid c) \qquad \text{[by 3]}$$
$$= [\,\{\,a' \mid (b \mid c)\,\}'\,]' \qquad \text{[by 3]}$$
$$= [\,(b' \mid a') \mid (c' \mid a')\,]' \qquad \text{[by 5, } a \text{ replaced by } a'\text{]}$$
$$= [\,(a' \mid b') \mid (a' \mid c')\,]' \qquad \text{[by } A\text{, used twice].}$$

Proof of V.

Take $\tilde{a} = a'$. Then

$$(a \mid a')' = (x \mid x')' \qquad \text{[by } B\text{, } b \text{ replaced by } x\text{]}$$
$$= z' \qquad \text{[by II}a\text{]}$$
$$= u \qquad \text{[by II}b\text{].}$$

$$a' \mid (a')' = a' \mid a \qquad \text{[by 3]}$$
$$= a \mid a' \qquad \text{[by } A\text{]}$$
$$= z \qquad \text{[by } B\text{, } b \text{ replaced by } x\text{].}$$

Postulates 1–5 *and Boolean Algebras.*

The following is Huntington's first postulate-set* for Boolean algebras:

"[For this postulate-set] we take as the fundamental concepts a class K with two [binary K-] rules of combination $\oplus$ and $\odot$; and as the fundamental propositions the following ten postulates:

Ia. $a \oplus b$ is in the class whenever a and b are in the class.

Ib. $a \odot b$ is in the class whenever a and b are in the class.

IIa. There is an element† z such that $a \oplus z = a$ for every element a.

IIb. There is an element† u such that $a \odot u = a$ for every element a.

IIIa. $a \oplus b = b \oplus a$ whenever $a, b, a \oplus b$, and $b \oplus a$ are in the class.

IIIb. $a \odot b = b \odot a$ whenever $a, b, a \odot b$, and $b \odot a$ are in the class.

IVa. $a \oplus (b \odot c) = (a \oplus b) \odot (a \oplus c)$ whenever $a, b, c, a \oplus b, a \oplus c, b \odot c, a \oplus (b \odot c)$, and $(a \oplus b) \odot (a \oplus c)$ are in the class.

IVb. $a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c)$ whenever $a, b, c, a \odot b, a \odot c, b \oplus c, a \odot (b \oplus c)$, and $(a \odot b) \oplus (a \odot c)$ are in the class.

V. If the elements z and u in postulates IIa and IIb exist and are unique, then for every element a there is an element $\bar{a}$ such that $a \oplus \bar{a} = u$ and $a \odot \bar{a} = z$.

VI. There are at least two elements, x and y, in the class such that $x \neq y$."‡

That set 1–5 is a postulate-set for Boolean algebras we shall prove by showing that this set and Huntington's first postulate-set are equivalent.

Proof.—If for any elements, a and b, of our class K we write

$$\bar{a} \text{ for } a', \quad a \oplus b \text{ for } (a \mid b)', \text{ and } a \odot b \text{ for } a' \mid b',$$

theorems Ia-V and postulate 1 are precisely Huntington's first postulate-set; hence set 1–5 implies Huntington's set.

If for any elements, a and b, of Huntington's class we write

$$a \mid b \quad \text{for} \quad \bar{a} \odot \bar{b},$$

Huntington's set implies set 1–5.§

§ 2. *Application to Primitive Logical Constants.*

Since not only in special deductive systems but even in the foundations of logic not all propositions can be proved and not all non-propositional entities

* HUNTINGTON, *loc. cit.*, pp. 292–3.

† For z and u respectively HUNTINGTON uses the symbols $\wedge$ and $\vee$, which he takes from PEANO's *Formulaire de Mathématiques*. These are, however, symbols for *logical constants*, just as 0 and 1 are symbols for *numerical constants*. We have replaced, therefore, BOOLE's and SCHRÖDER's 0 and 1, and HUNTINGTON's $\wedge$ and $\vee$, by z and u.

‡ That is, such that x and y are distinct.

§ By the "principle of duality" the results of § 1 hold also when $a|b$ is interpreted throughout as $\bar{a} \oplus \bar{b}$.

can be defined, some logical constants* must be primitive,† that is, either unproved or undefined. A list of primitive logical constants—primitive ideas and primitive propositions—in terms of which, presumably,‡ all other logical constants can be either defined or proved, is given by Whitehead and Russell in their *Principia Mathematica*.§ This list contains, among other logical constants, the primitive ideas‖ *negation* (symbolized $\sim$) and *disjunction* (symbolized $\vee$). Negation and disjunction are partly explained—but, of course, not at all defined—by the statement that for any elementary proposition p (*elementary proposition* being itself one of the *Principia's* primitive ideas), $\sim p$ means the elementary proposition *not-p*; and for any two elementary propositions, p and q, $p \vee q$ means the elementary proposition *either p or q* (*or both*).

On these two primitive ideas, in view of the following interpretation of K and $|$, our set 1–5 has an important bearing. For, if K is the class of all propositions of a given logical type,¶ then whenever p and q are two propositions of this type, $p \mid q$ may be interpreted as the proposition *neither p nor q*; in other words, $|$ has the properties of the logical constant *neither-nor*. This logical constant we may symbolize by $\wedge$, and for obvious reasons we may name *rejection*.**

THEOREM 1. *If in any list of primitive ideas for logic both negation and disjunction are primitive, they may be replaced by the single primitive idea rejection.*

Proof.—In terms of negation and disjunction, rejection is defined by the

Def.—For any two elementary propositions, p and q,

$$p \wedge q = \sim (p \vee q).$$

In terms of rejection, negation is defined †† by the

Def.—For any elementary proposition p, $\sim p = p \wedge p$.

In terms of rejection, disjunction is defined by the

Def.—For any two elementary propositions, p and q,

$$p \vee q = (p \wedge q) \wedge (p \wedge q).$$

By the following theorem, a similar reduction is possible for primitive propositions.

* WHITEHEAD and RUSSELL, *loc. cit.*, pp. 94–101.

† *Ib.*, p. 95.

‡ ". . . there must always be some element of doubt, since it is hard to be sure that one never uses some principle unconsciously" (*ib.*, p. 94).

§ WHITEHEAD and RUSSELL, *loc. cit.*, partial list, pp. 95–101; the other primitives are scattered throughout the rest of the book.

‖ *Ib.*, p. 97.

¶ *Ib.*, pp. 39–68.

** By analogy with *subject* and *object*, we may call $p \wedge q$ the *reject* of p and q.

†† Negation may thus be considered as a special case of rejection.

Two primitive propositions of the *Principia** are:

$*$ 1.7. If p is an elementary proposition, $\sim p$ is an elementary proposition.

$*$ 1.71. If p and q are elementary propositions, $p \vee q$ is an elementary proposition.

THEOREM 2. *If in any list of primitive propositions for logic both* $*$ 1.7 *and* $*$ 1.71 *are primitive, they may be replaced by the single primitive proposition.*

$*$ 1.7′. *If p and q are elementary propositions, $p \wedge q$ is an elementary proposition.*

Proof.—If in $*$ 1.7 we replace p by $p \vee q$, $*$ 1.7 and $*$ 1.71 imply $*$ 1.7′.

If in $*$ 1.7′ we replace q by p, $*$ 1.7′ implies $*$ 1.7.

If in $*$ 1.7′ we replace p by $p \wedge q$ and q by $p \wedge q$, $*$ 1.7′, used twice, implies $*$ 1.71.†

Thus we have made it possible to reduce, by one each, the number of primitive ideas and of primitive propositions used in the *Principia* for the foundation of logic.

CORNELL UNIVERSITY,
February, 1913.

* WHITEHEAD and RUSSELL, *loc. cit.*, p. 101.

† By the "principle of duality" the results of §2 hold also when $p \wedge q$ is interpreted throughout as the logical constant *either not-p or not-q*.

FORMAL MODULAR INVARIANTS WITH APPLICATION TO BINARY MODULAR COVARIANTS*

BY

MILDRED SANDERSON

§ 1. *Introduction.*

Consider a system of forms $f_1(x_1, \cdots, x_m), \cdots, f_k(x_1, \cdots, x_m)$, whose coefficients $a_1, \cdots, a_r$ (arranged in a definite order) are independent variables. Let $b_1, \cdots, b_r$ be the coefficients (taken in the same order) of the forms derived from $f_1, \cdots, f_k$ by applying a given linear transformation T with integral coefficients. Let $F(a_1, \cdots, a_r)$ be a polynomial with integral coefficients and

$$D(a_1, \cdots, a_r) = F(b_1, \cdots, b_r) - F(a_1, \cdots, a_r)$$

be the polynomial in $a_1, \cdots, a_r$ with integral coefficients which is obtained from $F(b) - F(a)$ upon replacing $b_1, \cdots, b_r$ by their expressions in terms of $a_1, \cdots, a_r$ and the coefficients of T. In case two or more of the terms in the initial expression for D had the same literal part $a_1^{e_1} \cdots a_r^{e_r}$, such terms are assumed to have been combined additively into a single term. Let p be a prime. According to the definition by HURWITZ† (stated for a single form in two variables), $F(a_1, \cdots, a_r)$ is an invariant of $f_1, \cdots, f_k$ modulo p with respect to the transformation T if $D(a_1, \cdots, a_r) \equiv 0 \pmod{p}$, identically in $a_1, \cdots, a_r$, viz., if the coefficient of each term of D is divisible by p. He gave an interesting example of such an invariant with respect to all of the transformations T with integral coefficients whose determinant is not divisible by p, but did not construct a theory of invariants modulo p.

As a generalization we may employ transformations T with coefficients in the Galois field $GF[p^n]$ composed of the p^n elements

$$c_0 + c_1 j + c_2 j^2 + \cdots + c_{n-1} j^{n-1} \quad (c_0, \cdots, c_{n-1} = 0, 1, \cdots, p-1),$$

where j is a root of a congruence of degree n irreducible modulo p. A polynomial $F(a_1, \cdots, a_r)$ with integral coefficients is a *formal* invariant of $f_1, \cdots, f_k$, in the field, under transformation T if‡ $D(a_1, \cdots, a_r)$ is identi-

* Presented to the Society (Chicago), March 21, 1913, under different title.

† Archiv der Mathematik und Physik, ser. 3, vol. 5 (1903), p. 22.

‡ In the spirit of KRONECKER'S use of polynomials in indeterminates $a_1, \cdots, a_r$ with coefficients in any domain of rationality, we use polynomials with coefficients in the $GF[p^n]$. Cf. J. KÖNIG, *Theorie der algebraischen Grössen*, Leipzig, 1903, p. 31; H. WEBER, Mathematische Annalen, vol. 43 (1893), p. 526.

cally zero in the field as to $a_1, \cdots, a_r$, i. e., if the coefficient of each term of D, when reduced to the form $c_0 + c_1 j + \cdots + c_{n-1} j^{n-1}$ by means of the congruence satisfied by j, has each c_i a multiple of p. If $F_0, F_1, \cdots$ are such formal invariants (with integral coefficients), then $F_0 + F_1 j + \cdots$ is a formal invariant. In this manner, or by direct extension of the previous definition, we have the concept of a formally invariant polynomial with coefficients in the $GF[p^n]$.

We pass to the entirely different concept of *modular* invariants, introduced by DICKSON.* The coefficients $a_1, \cdots, a_r$ of the forms are now undetermined elements of the $GF[p^n]$. A polynomial $F(a_1, \cdots, a_r)$ with coefficients in that field is called a modular invariant of the system of forms under any given group G of linear transformations with coefficients in the field if, for each transformation of G, $D(a_1, \cdots, a_r)$ is zero in the field. To apply this test we may first express† D as a polynomial $\delta(a_1, \cdots, a_r)$ in which the exponents are all less than p^n, and then require that δ shall be identically zero in the field as to $a_1, \cdots, a_r$. We thus see clearly just how the difference in the definitions of formal and modular invariants affects the actual computations. Dickson‡ has given a very simple and elegant theory of modular invariants. No theory has been developed for formal invariants. However, there exists between the two subjects an interesting and important relation, which I shall develop in what follows. I take this opportunity to express my gratitude to Professor Dickson for his interest and many helpful suggestions, in particular for the present formulation of this introductory section.

§ 2. *Statement and application of the fundamental theorem.*

To any modular invariant i of a system of forms under any group G of linear transformations with coefficients in the $GF[p^n]$, there corresponds a formal invariant I under G such that $I = i$ for all sets of values in the field of the coefficients of the system of forms.

This theorem enables us, as in the algebraic theory of invariants, to construct covariants of a system of binary forms in x and y from invariants of this system and an additional linear form whose coefficients are y and $-x$, if the invariants are made formally invariant with respect to x and y. Moreover, every invariant will yield a covariant. Some invariants will yield two or more independent covariants. For example,

$$f = a_0 x^2 + 2a_1 xy + a_2 y^2, \qquad a_0 x^4 + a_1 x^3 y + a_1 xy^3 + a_2 y^4$$

* These Transactions, vol. 8 (1907), p. 205.

† By use of Galois's generalization $a^{p^n} = a$ of Fermat's theorem.

‡ These Transactions, vol. 10 (1909), p. 123; American Journal of Mathematics, vol. 31 (1909), p. 337.

are two independent covariants modulo 3 of f, but as modular invariants of f and an additional linear form they are to be identified. The universal covariant

$$\begin{vmatrix} x & y \\ x^{p^n} & y^{p^n} \end{vmatrix},$$

on the other hand, could not be obtained as an invariant of a linear form, because it vanishes when x and y are in the $GF[p^n]$, as we suppose the coefficients of our forms to be. Whether or not all the covariants of a system of forms can be expressed as functions of this universal covariant and the invariants of this system and a linear form is a question as yet unanswered. In some special cases this has proved to be the case. In later sections will be shown a practical method for constructing covariants together with an illustration of the fact just mentioned. (See §§ 6, 7.)

In order to prove the above theorem, we shall first establish an interesting lemma.

§ 3. *Linear factors of a certain determinant D.*

Lemma. Let $a_1, \cdots, a_r$ $(r > 1)$ *be arbitrary variables, and* $g_1, \cdots, g_r$ *given elements of the* $GF[p^n]$, $g_r \neq 0$. *Then the determinant*

$$(1) \qquad N = \begin{vmatrix} a_1, & \cdots, & a_r \\ a_1^{p^n}, & \cdots, & a_r^{p^n} \\ \cdot & \cdots & \cdot \\ a_1^{p^{(r-1)n}}, & \cdots, & a_r^{p^{(r-1)n}} \end{vmatrix}$$

is divisible in the field by the determinant

$$(2) \qquad D = \begin{vmatrix} a_1, & \cdots, & a_r \\ a_1^{p^n}, & \cdots, & a_r^{p^n} \\ \cdot & \cdots & \cdot \\ a_1^{p^{(r-2)n}}, & \cdots, & a_r^{p^{(r-2)n}} \\ g_1, & \cdots, & g_r \end{vmatrix}$$

and the quotient $Q = N / D$ *has the properties*

$$(3) \qquad \begin{aligned} Q &\neq 0 \quad \text{if} \quad a_1 = g_1, \cdots, a_r = g_r, \\ Q &= 0 \quad \text{if} \quad a_1 = e_1, \cdots, a_r = e_r, \end{aligned}$$

where $e_1, \cdots, e_r$ *are elements of the field not proportional to* $g_1, \cdots, g_r$.

We know that N is an invariant and is equal to the following product:*

$$N = \prod_{k=1}^{r} \prod_{b} (a_k + b_{k+1} a_{k+1} + \cdots + b_r a_r), \tag{4}$$

where the second product extends over the $p^{(r-k)n}$ sets $b_{k+1}, \cdots, b_r$ obtained when each b ranges independently over the elements of the $GF[p^n]$.

Let us consider the determinant D. We can see at once that it is divisible by

$$a_1 - \frac{g_1}{g_r} a_r, \tag{5}$$

since this divides all the two-rowed minors formed from the first and rth columns. If we transform the variables by setting

$$a_1 = A_1 + \sum_{j=2}^{r} k_j A_j, \qquad a_j = A_j, \qquad (j = 2, \cdots, r), \tag{6}$$

and subtract from the elements of the first column of D the products of the elements of the jth column ($j = 2, \cdots, r$) by k_j, we may write D in the form

$$D = \begin{vmatrix} A_1, & A_2, & \cdots, & A_r \\ \cdots & \cdots & \cdots & \cdots \\ A_1^{p^{(r-2)n}}, & A_2^{p^{(r-2)n}}, & \cdots, & A_r^{p^{(r-2)n}} \\ g_1 - \sum_{j=2}^{r} k_j g_j, & g_2, & \cdots, & g_r \end{vmatrix}. \tag{7}$$

Thus D will be invariant under all transformations (6) for which

$$\sum_{j=2}^{r} k_j g_j = 0. \tag{8}$$

Since $g_r \neq 0$, the first $r-2$ k's may be chosen arbitrarily, hence in $p^{(r-2)n}$ ways, and k_r will be uniquely determined. If we transform (5) by all the transformations (6) which satisfy (8) we get $p^{(r-2)n}$ different linear factors, which upon replacing the A_i by the a_i are

$$a_1 + \sum_{j=2}^{r} k_j a_j - \frac{g_1}{g_r} a_r. \tag{9}$$

These must divide D since D was invariant under all such transformations. From (8) we see easily that all of these linear factors vanish for $a_1 = g_1, \cdots, a_r = g_r$.

Since D is also divisible by

$$a_2 - \frac{g_2}{g_r} a_r, \qquad \cdots, \qquad a_{r-1} - \frac{g_{r-1}}{g_r} a_r, \tag{10}$$

* L. E. Dickson, these Transactions, vol. 12 (1911), p. 76, and references there given.

by using transformations similar to (6), namely,

$$(11)\quad \begin{cases} a_1 = A_1, \quad a_2 = A_2 + \sum_{j=3}^{r} k_j A_j, \quad a_j = A_j \quad (j = 3, \cdots, r), \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_j = A_j \ (j = 1, \cdots, r-2), \quad a_{r-1} = A_{r-1} + k_r A_r, \quad a_r = A_r, \end{cases}$$

we find that D is divisible by respectively

$$p^{(r-3)n}, \quad \cdots, \quad p^n, \quad 1$$

new linear factors all of which vanish for $a_1 = g_1, \cdots, a_r = g_r$. These

$$p^{(r-2)n} + p^{(r-3)n} + \cdots + p^n + 1 = w$$

factors are all distinct. Hence D is divisible by their product. But w is exactly the degree of D, so that, apart from the factor $\pm g_r$, D is equal to this product. It may be written

$$(12)\qquad D = \pm g_r \prod_{k=1}^{r} \prod_{[b]} (a_k + b_{k+1} a_{k+1} + \cdots + b_r a_r),$$

where the second product extends over all values of the b_j in the $GF[p^n]$ such that the linear form vanishes for $a_1 = g_1, \cdots, a_r = g_r$.

Since all these factors are contained in N, Q is integral. Now D contains all linear forms with unity for the coefficient of the a_j of lowest subscript, which vanish for $a_i = g_i$ $(i = 1, \cdots, r)$. Hence, none of the factors of Q vanish for $a_i = g_i$, so that $Q \neq 0$ for $a_1 = g_1, \cdots, a_r = g_r$.

Some of the factors of the quotient must vanish for $a_1 = e_1, \cdots, a_r = e_r$, where $e_1, \cdots, e_r$ are not proportional to $g_1, \cdots, g_r$. For, there are w factors of N which do vanish, and unless these are the same as the factors of D, Q must vanish. But $r - 1$ of the factors of D, namely,

$$a_1 - \frac{g_1}{g_r} a_r, \qquad a_2 - \frac{g_2}{g_r} a_r, \qquad \cdots, \qquad a_{r-1} - \frac{g_{r-1}}{g_r} a_r,$$

are linearly independent, and hence when equated to zero uniquely determine the ratios of the a's for which D vanishes. Therefore, if Q does not vanish for $a_1 = e_1, \cdots, a_r = e_r$, then $e_1, \cdots, e_r$ must be proportional to $g_1, \cdots, g_r$.

§ 4. *Definition of the classes C_{ij}.*

Consider a system S of forms $f_1, \cdots, f_t$ in m variables with coefficients $a_1, \cdots, a_r$. Following the method of Professor L. E. Dickson,* we shall give the coefficients of S special values in the $GF[p^n]$, and divide the resulting

*These Transactions, vol. 10 (1909), p. 123; American Journal of Mathematics, vol. 31 (1909), p. 337.

set of systems into classes C_i, such that S' and S'' shall belong to the same class if and only if they are equivalent under G. If S' is a particular system of forms $f_1', \cdots, f_t'$, and if k is a constant, we shall say that the system of forms $kf_1', \cdots, kf_t'$ is a multiple of S', and we shall denote it by kS'. Let c_i be a subset of C_i such that if s is in c_i no multiple of s is in c_i, and such that every system S in C_i is a multiple of some system in c_i. Let ρ be a primitive root in the $GF[p^n]$. Suppose e_i to be the smallest exponent for which s' and $\rho^{e_i} s'$ are equivalent under G, where s' is any system in c_i. Then e_i is a factor of $p^n - 1$ and

$$s', \quad \rho^{e_i} s', \quad \cdots, \quad \rho^{(d_i-1)e_i} s' \qquad (d_i e_i = p^n - 1) \tag{13}$$

are all contained in C_i. If S'' is a system in C_i not in set (13), $\rho^{e_i} S''$ will be in C_i. For, S'' may be transformed into s', s' into $\rho^{e_i} s'$, $\rho^{e_i} s'$ into $\rho^{e_i} S''$. Moreover any system S''' in C_i is equal to ρ^{ke_i} times a system s''' in c_i. For we may write

$$S''' = \rho^{ke_i+l} s''' \qquad (0 \leqq l < e_i).$$

Since $\rho^{p^n-1-ke_i} S'''$ is in C_i, $\rho^l s'''$ is in C_i. But e_i is the smallest exponent for which this is possible. Hence $l = 0$. Therefore we may write

$$C_i = \sum_{k=1}^{d_i} c_i \rho^{ke_i},$$

where Σ denotes an aggregate, and $c_i \rho^{ke_i}$ is the set of systems obtained by multiplying each system in c_i by ρ^{ke_i}. If $e_i = 1$, C_i will contain all the multiples of the systems in c_i. In general there are e_i different classes

$$C_{il} = \sum_{k=1}^{d_i} c_i \rho^{ke_i+l} \qquad (l = 0, \cdots, e_i - 1) \tag{14}$$

formed on the subsets $c_i, c_i\rho, \cdots, c_i \rho^{e_i-1}$ respectively. Thus a complete list of the classes is given by (14), where $i = 1, \cdots, q$, and C_{00} which contains only the identically zero system.

§ 5. *Characteristic invariants.*

All invariants of the system S of forms can be expressed in terms of characteristic invariants* i_{kl} which have the value one for systems of the class C_{kl} and the value zero for systems in the remaining classes. To prove the theorem of § 2, it suffices to construct formal invariants I_{kl} which reduce for values of the coefficients in the $GF[p^n]$ to the characteristic invariants i_{kl}. For the construction of the I_{kl} it is convenient to employ the invariants

$$J_i = \sum_{g_1, \cdots, g_r}^{c_i} [Q(g_1, \cdots, g_r)]^{d_i} \qquad (i = 1, \cdots, q), \tag{15}$$

* L. E. Dickson, l. c. (preceding foot-note).

where the sum extends over the different sets of coefficients in c_i. When the variables are transformed under G, the coefficients $a_1, \cdots, a_r$ undergo an induced transformation which is also linear with coefficients in the $GF[p^n]$. J_i is a formal invariant since the numerators are invariant apart from a factor which is the d_ith power of the determinant of the induced transformation, and the denominators are permuted among themselves apart from the same factor.

Consider any particular denominator

$$\begin{vmatrix} a_1, & \cdots, & a_r \\ \cdot & \cdots & \cdot \\ a_1^{p^{(r-2)n}}, & \cdots, & a_r^{p^{(r-2)n}} \\ g_1', & \cdots, & g_r' \end{vmatrix}^{d_i}. \tag{16}$$

Since $g_1', \cdots, g_r'$ are the coefficients of a system of c_i, this system is also in C_{i0}. After transformation this will become

$$\begin{vmatrix} \sum_{j=1}^{r} c_{1j} A_j & \cdots, & \sum_{j=1}^{r} c_{rj} A_j \\ \cdot & \cdots & \cdot \\ \Sigma c_{1j} A_j^{p^{(r-2)n}}, & \cdots, & \Sigma c_{rj} A_j^{p^{(r-2)n}} \\ \Sigma c_{1j} G_j, & \cdots, & \Sigma c_{rj} G_j \end{vmatrix}^{d_i}, \tag{17}$$

where $G_1, \cdots, G_r$ are also the coefficients of a system in C_{i0} and hence may be written $\rho^{le_i} g_1'', \cdots, \rho^{le_i} g_r''$, where $g_1'', \cdots, g_r''$ are the coefficients of a system in c_i. Hence (17) becomes

$$\rho^{le_i d_i} \mid c_{tj} \mid^{d_i} \begin{vmatrix} A_1, & \cdots, & A_r \\ \cdot & \cdots & \cdot \\ A_1^{p^{(r-2)n}}, & \cdots, & A_r^{p^{(r-2)n}} \\ g_1'', & \cdots, & g_r'' \end{vmatrix}^{d_i}. \tag{18}$$

Since $\rho^{le_i d_i} = 1$, this denominator apart from the factor $\mid c_{tj} \mid^{d_i}$ occurs among the denominators in the sum (15), and the factor $\mid c_{tj} \mid^{d_i}$ is canceled by the same factor which is brought into the numerator by the transformation. Hence J_i is a formal invariant.

Let $J_i(C_{jk})$ denote the value which J_i has for systems of the class C_{jk}. Obviously $J_i(C_{00}) = 0$. In view of (15) and (3),

$$J_i(C_{kl}) = 0 \qquad (k \neq i), J_i(C_{il}) \neq 0. \tag{19}$$

Since $J_i(C_{i0}) \neq 0$, we may set

$$J_i(C_{i0}) = \rho^l. \tag{20}$$

We can easily see that

$$J_i(C_{ik}) = \rho^{l+kd_i} \qquad (k = 0, 1, \cdots, e_i - 1). \tag{21}$$

By making use of the fact that any invariant I may be written in the form*

$$I = \sum_{jk} v_{jk} I_{jk},$$

where v_{jk} is the value which I has for systems in the class C_{jk}, we obtain the following equations for the determination of the I_{ik}:

$$J_i^g = \sum_{k=0}^{e_i-1} \rho^{g(l+kd_i)} I_{ik} \qquad (g = 1, \cdots, e_i). \tag{22}$$

The determinant of the coefficients is

$$\left| \rho^{g(l+kd_i)} \right| = \rho^{l(1+2+\cdots+e_i)} \left| (\rho^{d_i})^{gk} \right| \quad (k = 0, \cdots, e_i - 1;\ g = 1, \cdots, e_i). \tag{23}$$

This is not zero since it is equal to the product of differences of distinct powers of ρ. Hence we can solve (22) uniquely for the I_{ik}.

The characteristic invariant for the class C_{00} is

$$I_{00} = 1 - \sum_{g_1, \cdots g_r}^{c} [Q(g_1, \cdots, g_r)]^{p^n - 1} \tag{24}$$

where c is the set of all subsets c_i.

§ 6. *Construction of binary covariants.*

Let us consider the covariants of a form f of the mth degree with coefficients in the $GF[p^n]$ under the group of all binary linear transformations with coefficients in the $GF[p^n]$. We may assume that it has been broken up into m linear factors whose coefficients are in a Galois Field of order $p^{m!n}$ or lower. We shall write it

$$f = \prod_{j=1}^{m} (\alpha_1^{(j)} x_1 + \alpha_2^{(j)} x_2) = \prod_{j=1}^{m} \alpha_x^{(j)} = \prod_{j=1}^{m} x_{\alpha(j)}. \tag{25}$$

When the variables are transformed, the coefficients of each linear form are transformed contragrediently to the variables. Any function F of the α's and the x's which is invariant and which is unchanged when the superscripts on any two sets of α's are interchanged will be a covariant of the form f. If we call any two of the linear factors of f α_x and β_x, F may obviously contain

* L. E. Dickson, l. c.

functions of the type

$$\alpha_x, \quad \alpha_x{}^{p^n} = \alpha_1 x_1^{p^n} + \alpha_2 x_2^{p^n}, \quad x_a{}^{p^n},$$

$$\begin{vmatrix} \alpha_1^{p^n} & \alpha_2^{p^n} \\ \alpha_1^{p^{kn}} & \alpha_2^{p^{kn}} \end{vmatrix} = (\alpha^{p^n} \quad \alpha^{p^{kn}}),$$

(26)

$$\begin{vmatrix} \alpha_1^{p^n} & \alpha_2^{p^n} \\ \beta_1^{p^{kn}} & \beta_2^{p^{kn}} \end{vmatrix} = (\alpha^{p^n} \quad \beta^{p^{kn}}),$$

$$\begin{vmatrix} x_1^{p^n} & x_2^{p^n} \\ x_1^{p^{kn}} & x_2^{p^{kn}} \end{vmatrix} = (x^{p^n} \quad x^{p^{kn}}).$$

It is not necessary that F should be integral in these functions, but if the covariant is to be integral, F must be integral in the α's and x's.

§ 7. *The covariants of a binary quadratic form in the* $GF[p^n]$ $(p > 2)$.

Consider the form

$$f = ax_1^2 + 2bx_1 x_2 + cx_2^2 = \alpha_x \beta_x \tag{27}$$

with coefficients in the $GF[p^n]$. We wish to find the covariants of this form under the group of all binary linear transformations with coefficients in the $GF[p^n]$. The coefficient of the highest power of x_1 is a seminvariant, that is, it is invariant under all the transformations of the type

$$x_1 = x_1' + \lambda x_2', \qquad x_2 = x_2'. \tag{28}$$

We shall first find the seminvariants of the form, and then construct covariants from these.

Under the group of transformations (28), the forms f are seen to belong to one of the following classes:

$$C_{\rho^i, d}, \qquad C_{\rho^i}, \qquad {}_{\rho^i}C, \qquad C_0 \tag{29}$$
$$(d = 0, \rho, \rho^2, \cdots, \rho^{p^n-1}; \; i = 1, \cdots, p^n - 1),$$

where ρ is a primitive root in the $GF[p^n]$. The canonical forms are respectively:

$$\rho^i (x_1^2 + dx_2^2), \qquad \rho^i x_1 x_2, \qquad \rho^i x_2^2, \qquad 0. \tag{30}$$

The classes are completely characterized by the seminvariants

$$\begin{aligned} D &= \tfrac{1}{4} (\alpha\beta)^2 = b^2 - ac, \\ a &= \alpha_1 \beta_1, \\ R &= (\alpha_1^{p^n-1} \alpha_2 - \alpha_2^{p^n}) (\beta_1^{p^n-1} \beta_2 - \beta_2^{p^n}), \\ B &= \tfrac{1}{2} \sum_2 (\alpha^{p^n} \beta) \beta_1^{p^n-1} = a^{p^n-1} b - b^{p^n}, \end{aligned} \tag{31}$$

as is seen by the accompanying table which gives the value that each seminvariant has for the respective classes. They also form a fundamental set.

(32)

	$C_{\rho^i, d}$	C_{ρ^i}	${}_{\rho^i}C$	C_0
D	$-\rho^{2i} d$	$\rho^{2i}/4$	0	0
a	ρ^i	0	0	0
R	$(-d$ square$)$ 0 $(-d$ not square$)$ $4\rho^i d$ $(-d$ zero$)$ 0	0	ρ^i	0
B	0	$-\rho^i/2$	0	0

All seminvariants must then be functions of these. However, there are some others which have very simple expressions, but which as functions of these four might be very complicated. They are

$$B_j = \tfrac{1}{2} \sum_2 (\alpha^{p^n} \beta)^j \beta_1^{j(p^n-1)},$$

(33)
$$S = \sum_2 \alpha_1^{2\tau p^n} \frac{(\beta\beta^{p^{2n}})}{(\beta\beta^{p^n})} \qquad \left(\tau = \frac{p^n - 1}{2}\right)$$

$$q = (a^\tau + c^\tau) \sum_{i=0}^{\tau} a^i c^i b^{2\tau-2i} - a^{3\tau} - c^{3\tau}.$$

The only invariants among these are D and q, and these form a fundamental system of invariants.*

By comparing the values which the different seminvariants have for the different classes according to tables (32) and

(34)

	$C_{\rho^i, d}$	C_{ρ^i}	${}_{\rho^i}C$	C_0
q	$(-1)^i [d^{2\tau} - 1]$	0	$(-1)^{i+1}$	0
B_j	$\binom{j \text{ even}}{-d \text{ square}} 2^j \rho^{ji} (-d)^{\frac{j}{2}}$ (otherwise) 0	$\frac{1}{2}(-1)^j \rho^{ji}$	0	0
S	$(-d$ square$)$ 2 $(-d$ not square$)$ 0 $(-d$ zero$)$ 2	1	0	0

we find various syzygies, among which are

(35)
$$D^\tau + a^{2\tau} = S + a^\tau q,$$

(36)
$$q = R^\tau a^{2\tau} + D^{2\tau} S^{2\tau} a^\tau - S^{2\tau} a^\tau - R^\tau.$$

*L. E. Dickson, American Journal of Mathematics, vol. 31 (1909) p. 109; these Transactions, vol. 10 (1909), p. 132.

Covariants are formed very easily from the seminvariants a, $B = B_1$, B_j, S and powers of these by replacing α_1 and β_1 by α_x and β_x or $\alpha_{x^{p^n}}$ and $\beta_{x^{p^n}}$, and by replacing $\alpha_1^{p^n}$ and $\beta_1^{p^n}$ by $x_{\alpha^{p^n}}$ and $x_{\beta^{p^n}}$, wherever the α_1 and β_1 do not occur in one of the invariant determinants.

Thus with a as a leader we get the two distinct covariants:

$$f = \alpha_x \beta_x, \qquad f_1 = \tfrac{1}{2} \sum_2 \alpha_{x^{p^n}} \beta_x. \tag{37}$$

Covariants led by B_j, where $j < p^n$, are

$$\tfrac{1}{2} \sum_2 (\alpha^{p^n} \beta)^j \beta_x^{j(p^n-1)}, \qquad \tfrac{1}{2} \sum_2 (\alpha^{p^n} \beta)^j x_{\beta^{p^n}} \beta_x^{(j-1)p^n-j}, \tag{38}$$

$$\cdots\cdots, \qquad \tfrac{1}{2} \sum_2 (\alpha^{p^n} \beta)^j x_{\beta^{p^n}}^{j-1} \beta_x^{p^n-j}.$$

The seminvariant S will lead the covariants:

$$\sum_2 x_{a^{p^n}}^k \alpha_x^{(2\tau-k)p^n} \frac{(\beta\beta^{p^{2n}})}{(\beta\beta^{p^n})} \qquad (k = 0, 1, \cdots, p^n - 1). \tag{39}$$

A simple method has not yet been found for constructing covariants with a leading coefficient R.

In addition to those already mentioned, should be noted the universal covariants*

$$L = (xx^{p^n}), \qquad Q = \frac{(xx^{p^{2n}})}{(xx^{p^n})}. \tag{40}$$

For the general Galois Field of order p^n, a complete system of covariants of the binary quadratic form has not yet been found. For the case $p^n = 3$ Professor L. E. Dickson† has found the complete system, but not their expressions in terms of the present symbols. The fundamental covariants are included among those mentioned above. They are

$$\begin{aligned} D &= (\alpha\beta)^2 = b^2 - ac, \\ q &= \frac{(\alpha^3\beta)^2 (\alpha\beta)(\alpha\alpha^3) - (\alpha\beta^3)(\alpha\alpha^9)}{(\alpha\alpha^3)^2} \\ &= ac^2 + a^2 c + b^2 c + b^2 a - a^3 - c^3, \\ f &= \alpha_x \beta_x, \\ f_1 &= \tfrac{1}{2} \sum_2 \alpha_{x^3} \beta_x = ax_1^4 + bx_1^3 x_2 + bx_1 x_2^3 + cx_2^4, \\ L &= (xx^3), \end{aligned} \tag{41}$$

* L. E. Dickson, these Transactions, vol. 12 (1911), p. 1, p. 75.

† These Transactions, vol. 14 (1913), pp. 306–310.

$$Q = \frac{(xx^9)}{(xx^3)} = x_1^6 + x_1^4 x_2^2 + x_1^2 x_2^4 + x_2^6,$$

$$C_1 = \tfrac{1}{2} \sum_2 (\alpha^3 \beta)\, \beta_x^2$$

$$= (a^2 b - b^3)\, x_1^2 + 2\,(ac^2 - a^2 c + b^2 c - b^2 a)\, x_1 x_2 + (b^3 - bc^2)\, x_2^2, \tag{41}$$

$$C_2 = \sum_2 \frac{(\alpha\alpha^9)}{(\alpha\alpha^3)} x_{\beta^3}^2 + \alpha_x \beta_x q$$

$$= (D + a^2)\, x_1^2 - 2\,(ab + bc)\, x_1 x_2 + (D + c^2)\, x_2^2.$$

It should be noticed that no covariant led by R occurs in this set. In this case it is expressible in terms of the others. In this case D, a, B, and q completely characterize the classes (24), but, for $p^n > 3$, q cannot be substituted for R in characterizing the classes.

The covariant C_2 is an interesting one because of the fact that the covariants of lowest order led by D and a^2 alone are respectively

$$DQ, \qquad f^2, \tag{42}$$

which are of order six and four respectively. One might expect that $D + a^2$ would lead no covariant of order less than twelve. One of this order is

$$DQ^2 + f^6. \tag{43}$$